

Projecting extreme heat-related mortality in Europe under climate change

Submitted by

Chun Kit Ho

to the University of Exeter as a thesis for the degree of

Doctor of Philosophy in Mathematics

April 2010

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Abstract

The assessment of health impacts of extreme hot weather under climate change is important for adaptation and mitigation actions. This thesis has developed techniques for estimating changes in heat-related mortality in Europe, with a focus on extreme daily mortality counts. The use of these techniques is illustrated through the projections of extreme elderly mortalities for London, UK and Budapest, Hungary from 2010 to 2099, using temperature projections from the perturbed physics ensemble of the regional climate model HadRM3.

The present-day relationship between daily number of deaths and temperatures at each location is modelled by Poisson generalized additive models. In order to account for possible discrepancies in climate model simulations, temperature projections from HadRM3 are calibrated by two approaches, bias correction and change factor. These are based on assumptions on the relationships in location, scale and shape between observed and modelled temperature distributions. In particular, a novel method using the Box-Cox transformation is developed to correct the bias in the upper tails of present-day simulated temperature distributions. Finally, future mortalities are projected by driving the mortality models with calibrated temperature projections.

Results of temperature calibration show that the two calibration approaches give substantially different estimates of future extreme temperatures. The estimates of 10-summer temperature return level by the two approaches differ by more than 4 °C over many parts of Europe in the period 2070 to 2099. For London and Budapest, the effect of this calibration uncertainty on extreme temperature projections is comparable to the effect of the uncertainty in climate model parameters which is estimated by the range of perturbed physics ensemble estimates. These two sources of uncertainties, together with the uncertainty in how the mortality-temperature relationship is modelled, contribute to large uncertainties in extreme mortality projections. Assuming constant elderly population in the future, the projected change in the 2-summer return level of number of daily elderly deaths in the period 2070 to 2099 relative to the the present-day ranges from –12% to +75% for London and from –16% to +22% for Budapest.

Contents

Acknowledgements	7
List of figures	8
List of tables	12
List of symbols	13
List of acronyms	16
1 Introduction	18
1.1 Motivation	18
1.2 Aim, scope and strategy	22
1.3 Plan of thesis	25
2 Background	26
2.1 Aim	26
2.2 Heatwaves in Europe	26
2.2.1 Historical trends in extreme daily temperatures	27
2.2.2 Physical processes leading to heatwaves	29
2.3 Brief overview of climate modelling	31
2.3.1 Global climate models	31
2.3.2 Predicting regional climate changes	32
2.3.3 Limitations of climate model predictions	35
2.3.4 Quantification of climate prediction uncertainties	38
2.4 Hot weather and human health	39
2.4.1 Human response to heat	39

2.4.2	Heat disorders and heat-related mortality	41
2.4.3	Adaptation to heat	42
2.5	Previous assessments of climate change impacts on heat-related mortality .	43
2.6	Summary	48
3	Data	50
3.1	Aim	50
3.2	Two focus cities: London and Budapest	50
3.2.1	Meteorological observation station data	50
3.2.2	Demographic data	51
3.3	European gridded temperature observations E-OBS	55
3.4	Regional climate model HadRM3	56
3.4.1	Description of model simulations	57
3.4.2	Perturbed physics ensemble	58
3.5	Summary	60
4	Statistical modelling of heat-related mortality	61
4.1	Aim	61
4.2	Modelling strategies	61
4.2.1	Model framework	62
4.2.2	Choice of environmental covariates	63
4.2.3	Measurement locations of environmental covariates	66
4.2.4	Modelling the functional dependence on covariates	69
4.3	Model evaluation methodologies	71
4.3.1	Model adequacy	71
4.3.2	Model specification	73
4.4	Model description	74
4.5	Model evaluation and comparison	77
4.5.1	Model with air temperature effect	77
4.5.2	The effect of humidity	82
4.5.3	Specification of the dependence of mortality on air temperature . . .	84

4.5.4	The use of gridded observed temperature as a covariate	89
4.6	Summary	90
5	Regional climate model temperature calibration	91
5.1	Aim	91
5.2	General approaches for calibration	92
5.2.1	Overview of approaches	92
5.2.2	Transfer functions for bias correction	95
5.2.3	Transfer functions for change factor	98
5.2.4	Comparison of approaches	99
5.3	Calibration of HadRM3 temperatures	101
5.3.1	Estimation of parameters	102
5.3.2	Comparison of temperature distributions	103
5.3.3	Estimation of return levels of extreme temperatures	104
5.4	Calibration results for London	105
5.4.1	Comparison of temperature distributions	105
5.4.2	Extreme temperatures for standard HadRM3 simulation	108
5.4.3	Extreme temperatures for ensemble members	112
5.5	Calibration results for Budapest	113
5.5.1	Comparison of temperature distributions	113
5.5.2	Extreme temperatures for standard HadRM3 simulation	115
5.5.3	Extreme temperatures for ensemble members	118
5.6	Spatially extended calibration over Europe	119
5.6.1	Comparison of present-day observed and simulated temperatures	119
5.6.2	Causes of HadRM3 biases for extreme temperatures	122
5.6.3	Comparison of HadRM3 temperatures in different time periods	124
5.6.4	Extreme temperatures calibrated by bias correction	126
5.6.5	Extreme temperatures calibrated by change factor	130
5.6.6	Uncertainties of calibrated extreme temperatures	131
5.6.7	Limitations of calibration methods	133
5.7	Summary	134

6	Extreme heat-related mortality projections	135
6.1	Aim	135
6.2	Projection methodologies	135
6.2.1	Assumptions on population and mortality annual cycles	136
6.2.2	Sensitivity analyses	137
6.2.3	Estimation of return levels of extreme mortality counts	143
6.3	Projections of future extreme summer mortalities	143
6.3.1	Projections for London	143
6.3.2	Projections for Budapest	145
6.4	Sensitivity analyses	145
6.4.1	Climate projections uncertainty	145
6.4.2	Mortality modelling uncertainty	150
6.4.3	Population uncertainty	151
6.5	Summary	152
7	Conclusions and future work	154
7.1	Summary of main findings	154
7.2	Wider applications	157
7.3	Future research directions	158
A	Population projections	160
B	Generalized linear models and generalized additive models	163
B.1	Exponential family distribution	163
B.2	Generalized linear models	164
B.3	Generalized additive models	166
C	Generalized Pareto distribution	168
	Bibliography	171

Acknowledgements

First of all, I would like to express my gratitude to my supervisors, Prof David Stephenson and Prof Mat Collins. I am very thankful for the guidance, support and invaluable insights from David since 2003 when he was my undergraduate tutor in Reading and throughout this challenging yet fascinating PhD project here at Exeter. I am especially thankful to him who offered to me the opportunity to be the first PhD student in Exeter Climate Systems (XCS). I am also very thankful for the advice and support from Mat who has introduced me to the world of climate modelling and answered many of my queries on this area.

This thesis has been benefited from discussions with a number of people, including Prof Trevor Bailey on mortality modelling, Dr Chris Ferro and Dr Simon Brown (UK Met Office) on the calibration of climate model projections. I would like to thank them for their invaluable advice. I would also like to thank Prof John Thurn who helped me settle down in the School during my first year.

My thanks also go to Prof Glenn McGregor (University of Auckland) and Ms Sari Kovats (London School of Hygiene & Tropical Medicine) for their help in obtaining the mortality data used in this study. Station meteorological observation data are provided by UK Met Office via the British Atmospheric Data Centre.

Finally, I am grateful for the funding on this project from the College of Engineering, Mathematics and Physical Sciences (CEMPS), and also from the UK Met Office through the Cooperative Award in Science and Engineering (CASE).

List of Figures

Figure 1.1	Maps of observed mean air temperature anomalies in June 2003, August 2003 and July 2006	19
Figure 1.2	Schematic diagram showing the mortality projection approach	23
Figure 2.1	Maps of orography and annual mean temperature fields of HadRM3 and HadCM3	34
Figure 2.2	Globally averaged surface air temperature change in 21 st century projected by 21 IPCC GCMs under two emission scenarios	36
Figure 2.3	Schematic representation of human heat balance model	40
Figure 2.4	Typical relationships between daily expected mortality and air temperature	42
Figure 2.5	Examples of mortality model fits by previous mortality projections	46
Figure 3.1	Time series of daily mean values of air temperature, vapour pressure and daily mean elderly mortality counts in London	52
Figure 3.2	Time series of daily mean values of air temperature, vapour pressure and daily mean elderly mortality counts in Budapest	53
Figure 3.3	Time series of past and future projected elderly population in London and Budapest	54
Figure 3.4	Maps of observed and HadRM3 simulated (forced by ERA-40) mean air temperature on 1 August 1995	56
Figure 3.5	Maps illustrating the spatial resolution of HadRM3 over UK and Hungary	57
Figure 3.6	Boxplots of summer T^o and T^g simulated by HadRM3 perturbed physics ensemble members from 1970 to 1999	60

Figure 4.1	Scatter plots of summer T_i^s against T_i^o for London and Budapest . . .	67
Figure 4.2	Return level plots of summer T_i^s and T_i^o for London and Budapest . .	68
Figure 4.3	Time series of observed and predicted mortality counts by models TS and BL	79
Figure 4.4	Exponential transformed estimated smoothing function of air temperature in models TS and TO	80
Figure 4.5	Residual diagnostic plots for model TS	81
Figure 4.6	Exponential transformed estimated smoothing function of vapour pressure in model TSH	83
Figure 4.7	Predicted elderly mortality as a function of daily mean station air temperature for models TS and PTS	86
Figure 4.8	Mean residual life plots for observed and predicted mortalities by models TS and PTS	87
Figure 4.9	Autocorrelation function of deviance residuals of model TS	89
Figure 5.1	Schematic diagrams showing two approaches to calibrate climate model projections	93
Figure 5.2	Illustration of distribution mapping by two calibration approaches . .	94
Figure 5.3	Quantile-quantile plot of G against O and plots of sample quantile difference between G and O for London	106
Figure 5.4	Quantile-quantile plot of G' (2040-69) against G and plots of sample quantile difference between G' and G for London	107
Figure 5.5	Return level plots of different temperature variables for London . . .	109
Figure 5.6	Estimates of 10-summer return level of present-day and future temperatures in London simulated by different PPE members of HadRM3	112
Figure 5.7	Quantile-quantile plot of G against O and plots of sample quantile difference between G and O for Budapest	114
Figure 5.8	Quantile-quantile plot of G' (2040-69) against G and plots of sample quantile difference between G' and G for Budapest	115
Figure 5.9	Return level plots of different temperature variables for Budapest . . .	116

Figure 5.10	Estimates of 10-summer return level of present-day and future temperatures in Budapest simulated by different PPE members of HadRM3	118
Figure 5.11	Maps showing sample mean, standard deviation and sample moment skewness of O , G and their difference or ratio	120
Figure 5.12	Maps showing differences in sample 0.95 quantile, 0.99 quantile and estimated 10-summer return level between G and O	121
Figure 5.13	Maps showing location-adjusted and location and scale-adjusted 0.99 quantile difference between G and O	122
Figure 5.14	Maps showing differences in sample mean, 0.99 quantile and moment skewness and ratio of standard deviation between T^a and O	123
Figure 5.15	Maps showing sample mean of G' in three future periods and their differences relative to G	125
Figure 5.16	Maps showing standard deviation of G' in three future periods and their ratios relative to G	126
Figure 5.17	Maps showing differences in sample moment skewness of G' in three future periods relative to G	126
Figure 5.18	Maps showing differences in sample 0.99 quantile of G' in three future periods relative to G	127
Figure 5.19	Maps showing estimates of power transformation parameters λ_O and λ_G	127
Figure 5.20	Maps showing bias in 10-summer level of \hat{O}'_b for the present-day	128
Figure 5.21	Maps showing differences between 10-summer level of G' for three different periods and O	129
Figure 5.22	Maps showing differences between the 10-summer level of \hat{O}' calibrated by different methods and that of O	130
Figure 5.23	Maps showing differences in 10-summer level and sample mean and ratio of standard deviation between \hat{O}'_b calibrated by BC-LSB and \hat{O}'_c calibrated by CF-LS	132
Figure 5.24	Scatter plot of 10-summer levels of \hat{O}'_c against \hat{O}'_b for 11 PPE members	133
Figure 6.1	Schematic diagram showing sources of uncertainties in mortality projections	138

Figure 6.2	Schematic representation of possible future changes in summer mortality-temperature relationship	142
Figure 6.3	Estimates of 2-summer temperature return level in London and Budapest by different PPE members of HadRM3	146
Figure 6.4	Estimates of 2-summer mortality return level in London and Budapest by different PPE members of HadRM3	147
Figure 6.5	Scatter plot comparing 2-summer mortality level for 11 PPE members projected by model TO using different calibration methods	149
Figure 6.6	Scatter plot comparing 2-summer mortality level for 11 PPE members projected using different mortality models	151
Figure A.1	Map showing the statistical regions of Hungary and the location of Budapest	161
Figure A.2	Observed and projected elderly population sizes in Budapest and the Közép-Magyarország region of Hungary	161
Figure A.3	Proportion of Budapest elderly population in the Közép-Magyarország region	162

List of Tables

Table 1.1	Estimates of excess number of deaths during the European heatwaves in 2003 and 2006	20
Table 4.1	Parameter estimates and R-squared statistics of model fits of T^g regressed on T^o for London and Budapest	67
Table 4.2	Mortality model evaluation statistics	82
Table 5.1	Comparison of two model calibration approaches	100
Table 5.2	Sample mean, standard deviation and 10-summer return level of calibrated temperatures for London	110
Table 5.3	Sample mean, standard deviation and 10-summer return level of calibrated temperatures for Budapest	117
Table 6.1	Observed and projected mean and 2-summer return level of temperature and mortalities for London	144
Table 6.2	Observed and projected mean and 2-summer return level of temperature and mortalities for Budapest	145
Table 6.3	Ensemble mean of 2-summer mortality return level estimates using different combinations of mortality models and temperature calibration methods	148
Table 6.4	Ensemble mean of mean and 2-summer return level of number of elderly deaths using constant population assumption and population projection data	152

List of symbols

	conditional on
\sim	distributed as
$[\cdot]$	integer part
$\stackrel{d}{=}$	equality in distribution
$A(\cdot)$	Box-Cox transformation function
$B(\cdot)$	transfer function for bias correction calibration approach
$C(\cdot)$	transfer function for change factor calibration approach
C_n	rate of convective heat loss from the human body
$[c_1, c_2]$	upper and lower bounds of confidence intervals
d	component of deviance
D	deviance
\mathbb{E}	expectation (population mean)
E_{re}	rate of heat loss from the human body through respiration
E_{sw}	rate of heat loss from the human body through evaporation of sweat
e^s	vapour pressure observed at meteorological stations (hPa)
e_i^s	daily mean value of e^s (hPa)
F_O	cumulative distribution function of O
F_G	cumulative distribution function of G
$F^{-1}(\cdot)$	inverse of a cumulative distribution function
G	climate model simulated variable
G_p	p quantile of G
$(\cdot)_i$	order of individual (daily) values in regression models
K	rate of heat loss through conduction from the human body

l	log-likelihood
l_p	penalized log-likelihood
M	metabolic heat production rate
m	basis dimension of smoothing functions in GAM
n_o	number of observations in a statistical model
n_p	degrees of freedom in a statistical model
n_r	residual degrees of freedom in a statistical model
n_u	number of observations above a threshold
N	normal distribution
O	observed climate variable
O'_b	calibrated future climate variable (O') by the bias correction approach
O'_c	calibrated future climate variable (O') by the change factor approach
$O_{(k)}$	k^{th} order statistic in the sample of O
O_p	p quantile of O
P_i	interpolated daily elderly population
$Pr(\cdot)$	probability of an event
Q	quantile of a Poisson distribution
$[q_1, q_2]$	lower and upper bounds of prediction intervals
R_n	rate of radiative heat loss from the skin
r_i	time (in years), e.g. [0, 10] (Budapest) [0, 11] (London)
r_m	return period
s	standard deviation
T^{ap}	apparent temperature
T_i^{ap}	daily mean value of T^{ap}
T_i^a	simulated daily mean air temperature by HadRM3 forced by ERA-40 ($^{\circ}\text{C}$)
T_m	return level of a temperature variable (e.g. O and G) ($^{\circ}\text{C}$)
T_i^o	gridded observed daily mean air temperature from E-OBS data set ($^{\circ}\text{C}$)
T_i^g	simulated daily mean air temperature by HadRM3 forced by HadCM3 ($^{\circ}\text{C}$)
T^s	air temperature observed at meteorological observation stations ($^{\circ}\text{C}$)
T_i^s	daily mean value of T^s ($^{\circ}\text{C}$)

u	threshold used in a GP distribution
v^s	wind speed observed at meteorological observation stations (m s^{-1})
W	rate of work done by the body
X	covariates (explanatory variables) in a GLM or GAM
\mathbf{X}	set of covariates
Y	daily number of elderly deaths (random variable)
y	observations of daily number of elderly deaths
y_+	new mortality observation
α_O	location parameter of distribution of O
α_G	location parameter of distribution of G
β	parameters in regression models
$\boldsymbol{\beta}$	set of parameters in regression models
β_O	scale parameter of distribution of O
β_G	scale parameter of distribution of G
ϵ	residuals of a normal linear model
ϵ^d	deviance residuals of a GLM or GAM
λ_O	power parameter of Box-Cox transformation on O
λ_G	power parameter of Box-Cox transformation on G
μ	mean elderly death rate
η	linear predictor of a GLM or GAM
ρ	Pearson product-moment correlation coefficient
θ	canonical parameter of a distribution from the exponential family
Φ	cumulative distribution function of standard normal distribution
ϕ	dispersion parameter of a distribution from the exponential family

Notes:

- In Chapter 5, $(\cdot)'$ is used to indicate future values (2010-99), while $(\tilde{\cdot})$ is used to indicate Box-Cox transformed variables. The symbols O , G , O' and G' refer to daily mean air temperatures in Sections 5.3 to 5.6.
- $(\hat{\cdot})$ represents an estimate.

List of acronyms

AR4	fourth assessment report of the Intergovernmental Panel on Climate Change
AT	apparent temperature
BC-L	bias correction in location
BC-LS	bias correction in location and scale
BC-LSB	bias correction in location and scale for Box-Cox transformed variables
c.d.f.	cumulative distribution function
CET	Central England Temperature
CF-L	change factor in location
CF-LS	change factor in location and scale
d.f.	degree(s) of freedom
ECMWF	European Centre of Medium Range Forecasts
e.d.f.	effective degree(s) of freedom
ERA-40	ECMWF re-analysis (from September 1957 to August 2002)
GAM	Generalized Additive Model
GCM	Global Climate Model
GEV	generalized extreme value
GLM	Generalized Linear Model
GP	generalized Pareto
HadCM3	Hadley Centre coupled ocean-atmosphere model version 3
HadSM3	slab ocean configuration of HadCM3
HadRM3	Hadley Centre regional climate model version 3
IPCC	Intergovernmental Panel on Climate Change

IQR	interquartile range
MME	multi-model ensemble
p.d.f.	probability density function
PHEWE	Assessment and prevention of acute health effects of weather conditions in Europe
PPE	perturbed physics ensemble
PRUDENCE	Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects
RCM	Regional Climate Model
s.d.	standard deviation
SRES	Special Report on Emission Scenarios
UBRE	un-biased risk estimator
UKCP09	UK Climate Projections (published in 2009)
w.r.t.	with respect to

Chapter 1

Introduction

1.1 Motivation

Extreme heat causes severe impacts on public health in many places worldwide. In the United States, about 8000 deaths from 1979 to 2003 are related to the exposure to excessive heat (Centers for Disease Control and Prevention, 2006). This figure is more than 4 times higher than the cumulative number of deaths caused by tropical cyclones affecting the country from 1979 to 2006 (Blake et al., 2007). In Asia, over 5000 heatwave-related deaths were estimated in India from 1998 to 2007 (Hales et al., 2003; Gosling et al., 2009a).

The health impacts of extreme heat are serious in Europe, as demonstrated by the catastrophic outcomes of the heatwaves in summer 2003. In June and early August, two heatwaves affected western and central parts of the continent. Figures 1.1(a) and (b) show the observed mean air temperature anomalies in June and August 2003 relative to the average in the corresponding months from 1971 to 2000 using the European gridded surface temperature data set E-OBS (Haylock et al., 2008; see Section 3.3 for details). Observed temperature anomalies were generally above 4 °C over continental Europe and exceeded 6 °C over parts of France and Italy. The total number of excess deaths (number of deaths exceeding the level expected without weather effects for the corresponding period of time; refer to Section 4.2.4 for details) during that summer was estimated to be over 35,000 (Confalonieri et al., 2007). The estimates for individual countries with large number of

excess deaths are listed in Table 1.1. One of the worst affected countries was France. In Paris, maximum air temperatures exceeded 35 °C for 9 consecutive days. The night-time minimum air temperatures of 25.5 °C on 11th and 13th August were the highest recorded since 1873 (Cohen et al., 2005). The sustained extreme heat resulted in a rapid increase in daily number of deaths, most of which were the elderly living in retirement homes (Vandentorren and Empereur-Bissonnet, 2005). With inadequate preparedness by the health-care services and the general public, the French public health system was under enormous pressure: hospitals were overwhelmed by the influx of patients; cemeteries were unable to handle the bodies of those deceased (Michelon et al., 2005).

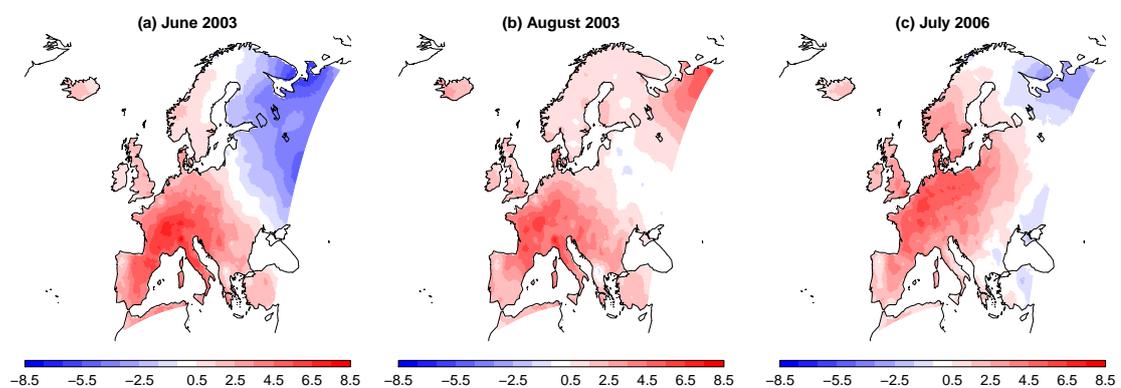


Figure 1.1: Observed mean air temperature anomalies (in °C) in (a) June 2003, (b) August 2003 and (c) July 2006 relative to the average in the corresponding months in 1971-2000.

Three years after, in July 2006, the western part of Europe was affected by another major heatwave. As shown in Table 1.1, the estimated number of excess deaths exceeded 2000 in France, but is considerably lower than that during the 2003 event. Fouillet et al. (2008) also estimates that this figure is less than one-third of the excess deaths predicted by a statistical model developed using historical mortality and weather data. The authors attributed the lower number of deaths to the improved awareness towards the risk of heatwaves and effective preventative actions which were put in place after the 2003 heatwave. However, in the Netherlands where the highest mean temperature anomalies (about 5 °C) were observed during the heatwave in July 2006 [Fig. 1.1(c)], the estimated number of excess deaths was 500 greater than that for the 2003 event. This suggests that extreme heat still poses a major threat to human health in Europe.

Year	Country	Period	Number of excess deaths	Increase from baseline	Reference
2003	France	1-20 August	14,802	60%	Pirard et al. (2005)
	Germany	1-24 August	1410	N/A	Kosatsky (2005)
	Italy	June - September	19,780	N/A	Kosatsky (2005)
	Netherlands	31 July - 13 August	500	3-5%	Garssen et al. (2005)
	Portugal	30 July - 12 August	1316	36%	Nogueira et al. (2005)
	Switzerland	June - August	975	7%	Grize et al. (2005)
	UK	4-13 August	2045	16%	Johnson et al. (2005)
2006	France	June - August	2090	9%	Fouillet et al. (2008)
	Netherlands	July	1000	N/A	Harmsen and Garssen (2006)
	UK	July	680	4%	UK Office of National Statistics (2006)

Table 1.1: Estimates of excess number of deaths during the European heatwaves in summers 2003 and 2006. Only countries where the estimates are 500 or greater are listed. Where available, the percentage increase from baseline mortality for the same period for each estimate are shown. Since the time periods covered and the methods of calculating the excess deaths vary among different studies, the estimates listed should be compared with caution.

Given the evidence of an increasing trend in the observed frequency of heatwave occurrences (see Section 2.2.1) and the projected climate change in the future, the health impacts of extreme heat have become a growing concern. According to the fourth assessment report of the Intergovernmental Panel on Climate Change Working Group I (IPCC AR4), global mean air temperatures are expected to rise by 1.1 °C to 6.4 °C by the end of the century compared to 1961 to 1990 (Meehl et al., 2007). In addition, the intensity, frequency and duration of heatwaves in Europe are likely (with a likelihood greater than 66%) to increase (Christensen et al., 2007b; Beniston et al., 2007). An assessment of the associated potential public health risk is required for planning appropriate response actions (Füssel et al., 2006). These include adaptation actions such as long-term investment in health care facilities and services, improvements to urban planning and building design standards (Kovats and Jendritzky, 2006), and mitigation actions which involve attempts to reduce greenhouse gases emission.

The assessment of future health impacts related to high temperatures, especially heat-related mortalities, has become an active area of research in recent years. Previous published mortality projections, for example Dessai (2002, 2003) and Gosling et al. (2007, 2009b), have adopted an approach which comprises of two components: climate modelling and health impact modelling (Haines et al., 2006; Gosling et al., 2009a). Meteorological variables in future summers under different greenhouse gases emission scenarios are projected by global or regional climate models. Health impact modelling involves using regression methods to empirically model the relationship between mortalities and meteorological variables based on past observations. Projections of future mortalities can then be made from the predictions of the regression model driven by climate model projections¹. As will be reviewed in Section 2.5, the techniques adopted by previous studies differ in several aspects, including how the health impact model is specified, which climate model is chosen and what sources of uncertainties are assessed. A common inadequacy

¹The terms ‘projection’ and ‘prediction’ carry different meanings in the climate science literature, including the IPCC AR4. Projections refer to the potential future evolution of a quantity based on certain assumptions, for example future greenhouse gases emissions, and are therefore subject to uncertainties. Predictions include estimates of such uncertainties, and are usually probabilistic in nature.

for these studies is that they have adopted simplistic methods to calibrate the possible discrepancies of climate model simulations.

Most of these studies have focused on the projected changes of the annual total heat-related mortalities. However, the events that happened in France during the 2003 heatwave demonstrate that changes in the frequency of days with very high number of deaths should be of particular concern. This is because the public health system is under the greatest pressure on these days, and long-term response actions should also consider projections of future *extreme* heat-related mortalities (McGregor et al., 2007).

1.2 Aim, scope and strategy

The main aim of this thesis is to develop techniques to project changes in extreme heat-related mortality in Europe under climate change. These techniques are illustrated through mortality projections from 2010 to 2099 using the output from perturbed physics ensemble of Hadley Centre regional climate model (RCM) HadRM3. The term ‘extreme heat-related mortality’ needs to be first defined. As will be described in Section 2.4, only a small proportion of the increased number of deaths on days with elevated temperatures is directly attributed to acute heat illnesses such as heat stroke (Kovats and Hajat, 2008). Therefore in this thesis, ‘heat-related mortality’ refers to the increased number of deaths from all causes related to increased air temperatures during the summer, as in AR4 of IPCC Working Group II (Section 8.4 in Confalonieri et al., 2007). As for the word ‘extreme’, which has various meanings in the climate science literature (Beniston et al., 2007; Stephenson, 2008), it is used in this thesis to refer to events with a low probability of occurrence (i.e. rare events) with a significant impact on the society. A day with ‘extreme heat-related mortality’ is one with exceptionally large number of deaths caused by high air temperatures in the summer. For mortality projection purposes in this thesis, summer is defined to be the four-month period from 15th May to 15th September, longer than the typical climatological definition (June to August). This choice is based on the observed relationship between mortality and temperature (see Section 4.4).

In order to make progress this thesis exploits the following two characteristics of heat-

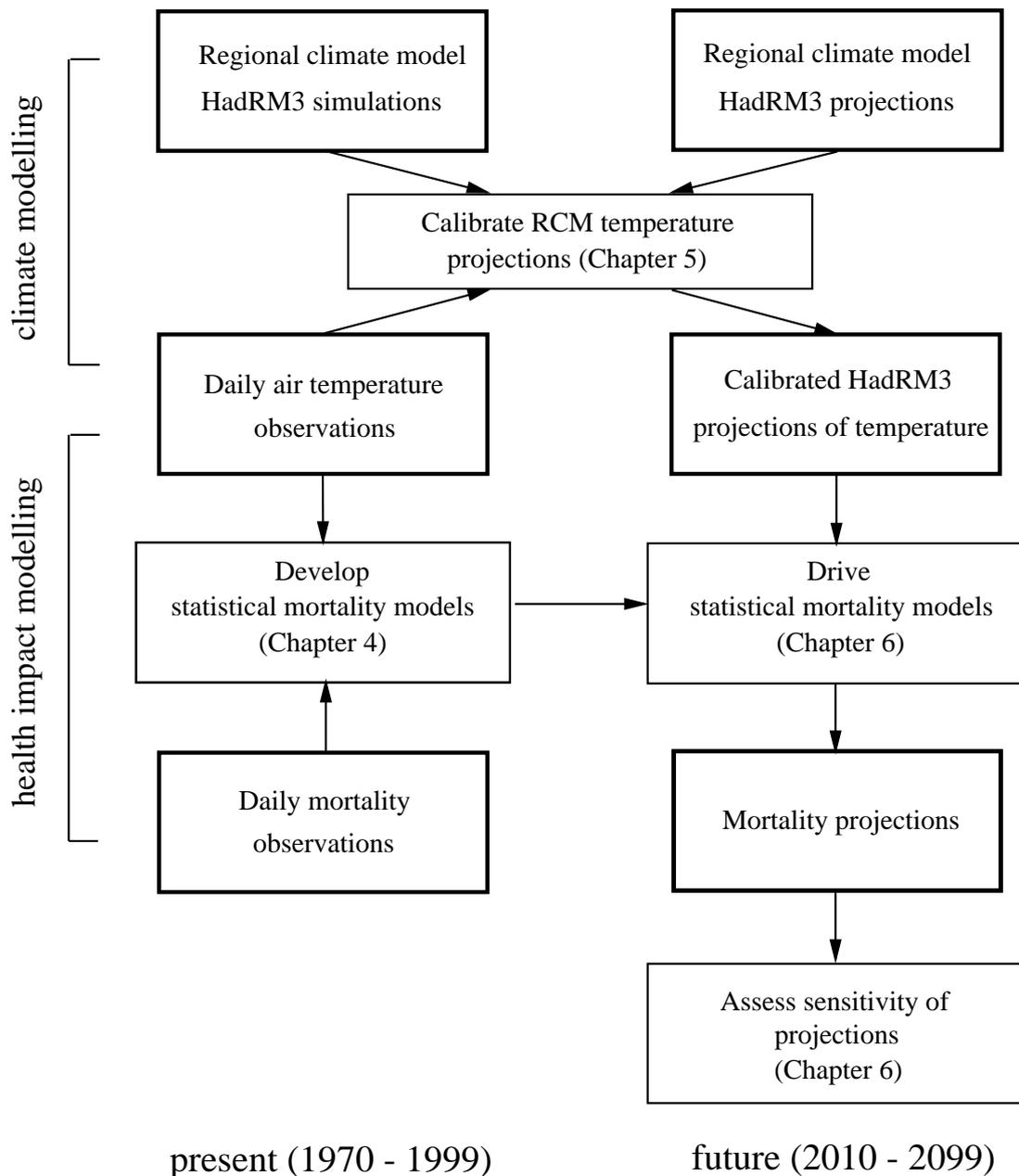


Figure 1.2: Schematic diagram showing the approach adopted in this thesis to project future heat-related mortality in London and Budapest. The chapters where the work of each step is described are also shown. Note that the ‘present’ period is defined to be 1970 to 1999 for RCM temperature calibration purposes. Other definitions are used for mortality model fitting - see Chapter 4 for details.

related mortality (see Section 2.4 for details). First, the elderly is the age group most vulnerable to extreme heat. The assessment and projections of mortality in this thesis therefore only consider this age group, defined as age 65 or above. Second, the relationship between mortality and temperature is location-specific. The projections of future heat-related mortality for the whole of Europe then require historical mortality data from numerous locations. In view of data availability and the focus of this thesis which is the illustration of mortality projection techniques rather than the projection results for each individual location in Europe, mortality projections are only assessed for two cities: London, UK and Budapest, Hungary. These two locations are chosen because of their different climate characteristics and sensitivity of mortality to high temperatures. For other locations in Europe within the domain of HadRM3 [see Fig. 3.4(b)], the calibrated projections of extreme temperatures from the standard version of HadRM3 will be presented. The corresponding impacts on heat-related mortalities could then be inferred by considering these projections together with the mortality-temperature relationships estimated for different European cities in other published studies, for example Keatinge et al. (2000) and McMichael et al. (2008).

As shown schematically in Fig. 1.2, the general approach to project future heat-related mortality for London and Budapest in this thesis is largely similar to that adopted by previous studies mentioned in the last section. There are, however, some differences in the techniques used, especially for the calibration of climate model projections. Statistical mortality models are first developed separately for the two cities using mortality and weather observation data, based on careful evaluation and comparison of possible modelling strategies. It will be concluded in Chapter 4 that daily mean air temperature is the only meteorological variable to be included in the models for mortality projection purposes. As a result, only the projections of air temperature from the 11 HadRM3 ensemble members are required. Using observed daily temperatures and HadRM3 simulated daily temperatures for the period 1970 to 1999, comprehensive calibration is performed on the raw HadRM3 temperature projections from 2010 to 2099 where different possible methods are tested. Finally, future changes of heat-related mortality are projected by driving the statistical mortality models with the calibrated HadRM3 temperature projections. The

sensitivity of mortality projections to different sources of uncertainties is also analysed.

1.3 Plan of thesis

The remainder of this thesis is structured as follows. Chapter 2 provides background information on extreme hot weather in Europe, its impact of human health and climate predictions. Previous studies on heat-related mortality projections are also reviewed. Chapter 3 describes the demographic and meteorological data sets and climate models used in this thesis. The following three chapters present the work of the three key steps in projecting heat-related mortalities shown in Fig. 1.2. Chapter 4 covers the specification, fitting and evaluation of the statistical mortality models for London and Budapest. Chapter 5 describes the approaches for calibrating temperature projections from HadRM3 and examines the future changes in the extreme calibrated temperatures over Europe. Chapter 6 presents the projections of heat-related mortalities in London and Budapest based on the calibrated temperature projections. A sensitivity analysis on uncertainties is also included. Finally, concluding remarks with a summary of main findings and possible future directions for research are given in Chapter 7.

Chapter 2

Background

2.1 Aim

This chapter briefly reviews the literature on three subject areas related to this thesis: extreme hot weather in Europe (Section 2.2), climate predictions using climate models (Section 2.3) and the impacts of extreme heat on human health (Section 2.4). In addition, previous studies on the projections of future heat-related mortality are reviewed (Section 2.5).

2.2 Heatwaves in Europe

The word ‘heatwave’ generally refers to periods of exceptionally warm temperatures, but there is no universal technical definition (Robinson, 2001). This is possibly because of the complex nature of these events. Like other extreme weather events, the impacts of heatwaves on the society (not only limited to human health) are determined by more than one of their attributes. These include their intensity, frequency and temporal duration (Stephenson, 2008). For intensity, the impacts of heatwaves do not necessarily depend on the absolute values of temperatures (or other related meteorological variables) as the society is able to adapt to local variations of climate to a certain extent. An example of such adaptation is ‘acclimatization’ to be described in Section 2.4.3. For certain impacts

of heatwaves, the rarity of extremely high temperatures may be more relevant.

This section first reviews studies which analysed the observed trends of extreme warm temperature events in Europe over the past century. The underlying physical processes of European heatwaves are then described. An understanding of these processes is important when interpreting outputs from the RCM in Chapter 5.

2.2.1 Historical trends in extreme daily temperatures

The availability of global and European daily observed temperature data has improved over the past decade. This has allowed more comprehensive studies of long-term changes in extreme daily temperatures (Alexander et al., 2006). Many of these studies examined such changes by analysing the trends of a set of standard indices (e.g. Frich et al., 2002; Klein Tank et al., 2009) which are related to different attributes of extreme temperature events. Klein Tank and Können (2003) analysed the trends of several extreme temperature indices for 86 European observation stations from 1946 to 1999. Using the Student's t-test, statistically significant (at 5% level) positive linear trends were detected for the index 'warm nights' in more than 75% of the stations for the period 1976 to 1999. This index is defined as the annual number of days where the minimum temperature exceeds the 90th percentile of daily minimum temperature for the corresponding calendar day in the reference period 1961 to 1990 [refer to the appendix in Klein Tank et al. (2009) for the formal definition]. Moberg and Jones (2005) considered different indices on the same temperature data set from 1901 to 1999, including the 90th percentile of daily minimum temperature and 90th percentile of daily maximum temperature for each summer (June to August). Statistically significant (at 5% level) positive linear trends for these indices were detected in 70% and 36% of the stations respectively. The above-mentioned 3 indices are related to the frequency and intensity of extremely warm temperatures. As for the duration of extremely warm temperatures, Alexander et al. (2006) calculated the trend of the gridded 'warm spell duration index' globally for the period 1951 to 2003. This index is defined as the annual count of days in a span of at least 6 days where the maximum temperature exceeds the 90th percentile of daily maximum temperature for the corresponding

calendar day in the reference period 1961 to 1990 (Klein Tank et al., 2009). Significant increase in this index was observed over parts of eastern Europe. The results from these studies provide some evidence that the frequency, intensity and duration of extremely warm temperatures in Europe have increased over the past century.

The use of percentiles in these extreme indices allows comparison of results from different geographical locations with different climate features because it focuses on the same part of the temperature distribution. However, a major limitation of these extreme indices is that they only describe ‘moderately extreme’ events which on average occur several times a year. These indices therefore do not represent changes of rarer extreme events which cause serious impacts (Klein Tank and Können, 2003). For example, the choice of the 90th percentile for the ‘warm nights’ index corresponds to a return period of only 36 days. This contrasts with the estimated return period of 46,000 years for the temperature recorded in summer 2003 in Switzerland (Schär et al., 2004).

Other methods can be applied to evaluate changes in rarer extremes of observed temperatures farther out in the tails of their distributions. Ferro et al. (2005) describes how distributions of weather variables can be compared for large gridded data sets, based on the differences in the quantiles of the distributions. This method can be used to assess the changes in extreme temperatures by considering quantiles in the tails of their distributions and the associated sampling uncertainties. Chapter 5 of this thesis will apply this method to analyse the temperature simulations of HadRM3. In addition, extreme value theory can also be used by incorporating time trends as covariates into the parameters of an appropriate extreme value distribution (Chapter 4 in Klein Tank et al., 2009), such as generalized extreme value (GEV) distribution for changes in temperature maxima and generalized Pareto (GP) distribution for changes in temperature excesses above thresholds (Chapter 6 in Coles, 2001). This is illustrated in Coelho et al. (2007) which fitted the GP distribution to historical gridded summer monthly temperatures in the northern hemisphere. However, among the literature reviewed for this thesis, there were no published studies using these two methods to examine historical changes in observed extreme daily temperatures for station-based or gridded data sets on a continent-wide or global basis.

2.2.2 Physical processes leading to heatwaves

Heatwaves in Europe are often related to a large-scale circulation pattern known as ‘blocking’ (Cassou et al., 2005; Carril et al., 2008). During blocking episodes, an anticyclone (area of high pressure) remains quasi-stationary over continental Europe for days or even weeks, while the normal eastward progression of low pressure systems is disrupted (Kyselý, 2008; Trigo et al., 2005). Rainfall within the anticyclone is suppressed by vertical subsidence of air. Prolonged clear-sky conditions allow intense heating of the ground by solar radiation at daytime and therefore lead to high surface temperatures (Garcia-Herrera et al., 2005; Grazzini et al., 2003). Night-time cooling is limited by the release of heat accumulated on the ground during the day (Black et al., 2004), keeping temperatures elevated for a long period of time.

The interactions between the land and the atmosphere also contribute to the occurrence of heatwaves. The evaporation of soil moisture near the surface cools the surrounding air and thus offsets the rise in air temperatures. This means that heatwaves will be more intense if this evaporative cooling diminishes because of decreased soil moisture (Clark et al., 2006). Wetherald and Manabe (1995, 1999) proposed two (related) mechanisms which can cause the drying of soil in the mid-latitudes in the summer. First, soil moisture is reduced in the antecedent spring due to enhanced evaporation or a deficit in precipitation. With drier soil, evaporation of moisture in the summer is reduced, leading to a reduction in rainfall. Second, this reduction in rainfall dries the soil in the summer, which in turn reduces rainfall further. A positive feedback of soil drying is then developed (see also Rowell and Jones, 2006).

A number of studies have examined the relative roles of the circulation pattern and land-atmosphere interactions in past European heatwaves. In a study of large-scale forcing factors which contribute to heatwaves in Europe, Della-Marta et al. (2007) performed a canonical correlation analysis¹ on surface temperature and precipitation data from Europe and sea level pressure and sea surface temperature data from Europe and North At-

¹The canonical correlation analysis technique relates two multivariate data sets by identifying linear combinations of variables in each data set, such that the correlation between the new variables is maximized (Chapter 12 in Wilks, 2006).

lantic for the period 1880 to 2005. It was found that heatwaves over western Europe are most strongly associated with positive sea level pressure anomalies over the Scandinavian region. Fischer et al. (2007a) attempted to quantify the effect of land-atmosphere interactions on the severity of four past European heatwaves from 1976 to 2005 using simulations on a RCM. Such interactions were found to account for 50 to 80% of the number of hot summer days. The contributions of circulation patterns and land-atmosphere interactions to European heatwaves are not totally independent. For the exceptional 2003 heatwave, anticyclonic conditions dominated central Europe in spring and summer of the year (Beniston and Diaz, 2004; Fink et al., 2004), leading to an early depletion in soil moisture over the region from spring onwards (Fischer et al., 2007b). The soil moisture feedback then exacerbated the rise of temperatures during the summer. Fischer et al. (2007b) suggests that this increase in temperatures further amplified the pre-existing anticyclonic circulation, which again led to temperature and soil moisture anomalies.

Some simulations of global climate models, such as Gregory et al. (1997) and Wetherald and Manabe (2002), project drying summers with reduced soil moisture over continental Europe under anthropogenic climate change. The role of soil moisture in driving extreme temperatures is therefore expected to become more important in the future. Fischer and Schär (2009) analysed the summer temperature projections by 11 different RCMs and decomposed the future changes in temperature variance into interannual, intraseasonal and seasonal components. Among the different RCMs, interannual temperature variance was found to increase by 30 to 95% over France by the end of the 21st century relative to 1961 to 1990. Fischer and Schär (2009) suggested that this change is mainly related to the growing importance of land-atmosphere interactions, and this increase in variance could lead to a higher intensity of temperature extremes in the future. Nevertheless, Rowell and Jones (2006) notes that because of uncertainties in climate model formulation and future anthropogenic emissions (see Section 2.3.3), there are still considerable uncertainties in the magnitude of projected summer drying caused by large-scale warming and soil moisture depletion in spring. There are also uncertainties in future changes in the frequency and strength of atmospheric blocking, as most climate models have shown limited ability in simulating these events (D'Andrea et al., 1998; Annexes 3 and 6 in Murphy et al.,

2009).

2.3 Brief overview of climate modelling

This section gives a brief overview of how climate models are used to produce global and regional climate projections for the future. Sources of biases and uncertainties in climate predictions will then be discussed. The existence of biases requires the calibration of RCM projections, while the uncertainties in climate projections contribute to the overall uncertainties in the heat-related mortality projections.

2.3.1 Global climate models

Global climate models (GCMs) (often referred to as general circulation models) are the most sophisticated class of climate models. The Earth's physical climate system includes the atmosphere, the ocean, the land (land surface and soil) and the cryosphere (ice on land and over the seas). These models represent the large-scale flows, processes, interactions and feedbacks within and among different components of the climate system by mathematical equations based on physical laws (McGuffie and Henderson-Sellers, 2005). In GCMs, the atmosphere and oceans are represented by values at points on three-dimensional grids. At each grid-point, the time-dependent mathematical equations are solved at each time step using numerical techniques.

The current generation of GCMs generally represent the atmosphere and the ocean with different configurations, primarily because processes within these two components of the climate system operate at different length and time scales. The two components are then 'coupled' at regular intervals. The resolutions of GCMs are limited by the computing resources available. For the 23 GCMs widely used in AR4 of IPCC (Randall et al., 2007), the atmosphere components generally have horizontal resolutions ranging from 1° to 4° of latitude and longitude (about 120 to 500 km) and around 20 vertical levels. For the ocean components, the horizontal resolutions generally range from 0.5° to 3° (about 60 to 350 km) and there are about 30 vertical levels. Many processes take place at a scale

smaller than the spatial and temporal resolution of GCMs, for example convective cloud formation. These have to be represented by the relationship between the model-resolved variable and the spatially or temporally averaged effects of such processes (Chapter 5 in McGuffie and Henderson-Sellers, 2005). Such a technique is known as parameterization.

Changes in the concentration of greenhouse gases (such as carbon dioxide and methane) and aerosols can affect the climate by altering the radiative balance of the earth-atmosphere system. Increases in anthropogenic emissions of greenhouse gases have very likely caused the observed widespread warming in the past few decades, despite an offset by aerosols (Hegerl et al., 2007; Stott et al., 2010). The future concentration of these forcing agents is therefore important in climate predictions. Many GCM simulations for climate projections, including the one used in this thesis (HadCM3, see Section 3.4), are driven by greenhouse gases and sulphur concentration under 6 groups of emission scenarios developed by IPCC (Nakićenović and Swart, 2000). These ‘SRES scenarios’ are based on plausible storylines of future changes in demographic, socioeconomic, technological and environmental developments. They cover situations from a future world which uses fossil fuel intensively, to one where alternative energy sources are dominant. However, they do not assume any climate change mitigation policies aiming to reduce emissions. Probabilities are not assigned to the different scenarios.

2.3.2 Predicting regional climate changes

Climate change impact assessments, such as temperature-related mortality, crop yields and flood risk, often require climate predictions at a local or regional scale. There has been an improvement in the horizontal resolution of GCMs in recent years (Randall et al., 2007). However, the detailed spatial structure of climate variables, for example temperature and precipitation, cannot yet be resolved by these models, especially around coastlines and over regions with complex topography or land surface distribution (Christensen et al., 2007a). Two approaches, statistical downscaling and dynamical downscaling, have been commonly adopted to produce regional-scale projections from the output of GCMs.

Statistical downscaling involves the construction of a statistical model where the local

surface variable of interest is projected using large-scale atmospheric variables modelled by the GCM, such as sea-level pressure, surface and upper-level temperature and humidity (Wilby and Wigley, 1997). This approach is often applied to project future precipitation for a point because the high spatial and temporal variability of this variable cannot be adequately represented by GCMs (Leung et al., 2003). Nevertheless, many studies also apply this approach to downscale other variables. For example, Hayhoe et al. (2004) projected future heat-related mortalities in California, USA, based on statistically downscaled temperature projections from GCMs under different emission scenarios. Various statistical downscaling techniques are available. Regression models can be used to establish the linear or non-linear relationships between local surface variable and large-scale variables for the present time. Given the projections of large-scale variables for the future from the GCM, the local variable can be predicted using the regression model. For example, Murphy (1999) used a multiple linear regression model to predict temperature and precipitation for 976 European stations, using surface wind velocity, vorticity, upper-level temperature and humidity from a GCM with horizontal resolution of about 300 km as explanatory variables. On the other hand, weather generators can produce realistic synthetic sequences of a local surface variable for the future (Wilks and Wilby, 1999). The parameters of weather generators, which describe the statistical properties of the local variable (such as mean, variance and auto-correlation), are conditioned on the large-scale variables from the GCM. The relative merits and limitations of different statistical downscaling techniques are discussed in Wilby et al. (2004). Regardless of which technique is used, the statistical downscaling approach assumes that the relationship between the local variable and large-scale variables from the GCM do not change with time and that the chosen large-scale variables fully represent the climate change signal (Giorgi et al., 2001).

The temperature projections used in this thesis are based on another downscaling approach called dynamical downscaling. This approach uses a limited area model or a regional climate model (RCM) nested within a GCM to generate high-resolution climate projections over a smaller spatial (and in some applications, temporal) domain, typically 5000 km × 5000 km. The boundary conditions of the RCM, including the large-scale atmospheric fields, sea-surface temperature, sea ice extent and thickness, are provided by

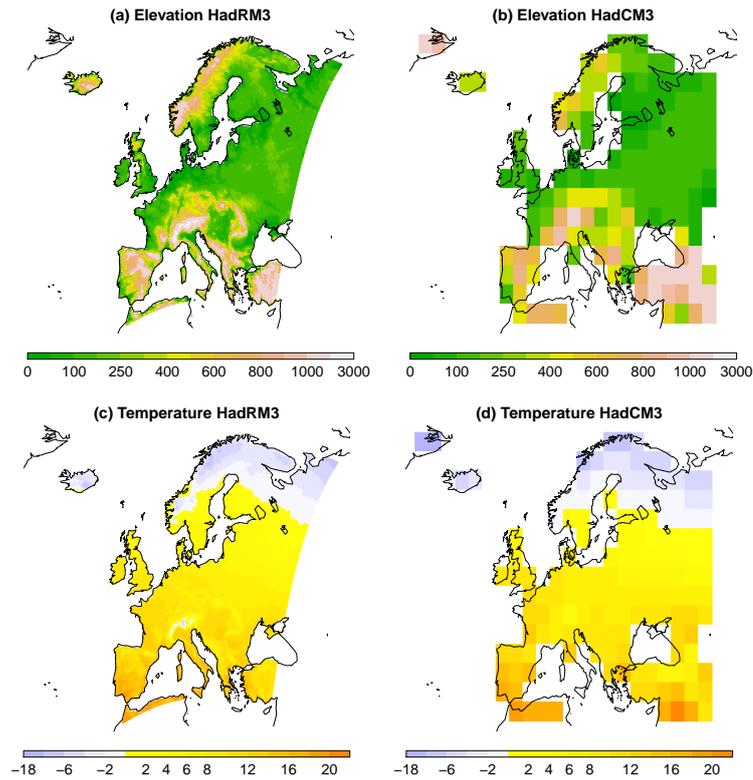


Figure 2.1: Model orography and temperature fields of the RCM HadRM3 used in this thesis and its driving GCM HadCM3 for Europe. Panels (a) and (b) show the elevation from sea level (in m). Panels (c) and (d) are the simulated annual mean surface temperatures (in °C) from 1961 to 1990.

a GCM at regular intervals. However in general, such ‘nesting’ is one-way, i.e. the results from RCM simulations do not affect the simulations of the driving GCM (Giorgi et al., 2001; Christensen et al., 2007b). The horizontal grid spacing of a RCM is typically 50 km or less. Figures 2.1(a) and (b) shows the model orography fields of HadRM3 and HadCM3 for Europe. These are the RCM used in this thesis (minimum grid spacing of about 25 km) and its driving GCM HadCM3 (grid spacing of about 300 km at the mid-latitudes) respectively (see Section 3.4 for details). It is clear that HadRM3 can much better represent the coastal regions and topography over Europe, for example Italy, the Alps and the Scandinavian Mountains. The higher resolution of RCMs enables certain features to be better simulated compared to GCMs, including regional details of the climate, extreme weather events and mesoscale weather systems, such as tropical and extratropical cyclones (Jones

et al., 2004). This can be demonstrated by Figs 2.1(c) and (d) which show the simulated annual mean temperatures from 1961 to 1990 by HadRM3 and HadCM3 respectively. The former can provide more detailed spatial distribution of temperatures. Despite the increased resolution, parameterization is still necessary in RCMs to represent sub-gridscale physical processes.

Comparing the above two downscaling approaches, dynamical downscaling is more computationally expensive, but projections using this method are from physically based models and are in general consistent with the large-scale projections from the driving GCM (Jacob et al., 2007; Mearns et al., 2003). These two approaches can complement each other for certain applications. For example, UK Climate Projections (UKCP09) used the weather generator technique to produce future sub-daily time series of temperature, vapour pressure and sunshine duration at a $5 \text{ km} \times 5 \text{ km}$ grid based on the simulations from HadRM3 (Jones et al., 2009).

2.3.3 Limitations of climate model predictions

Predictions from global and regional climate models are subject to uncertainties and biases from different sources.

The first source of uncertainty is related to climate forcing agents (emissions uncertainty or scenario uncertainty). Future anthropogenic emissions of greenhouse gases and aerosols depend on a number of socio-economic factors. This uncertainty can be accounted for by driving climate models with different SRES scenarios mentioned in Section 2.3.1, as has been done in the projections of future global mean temperatures for AR4 of IPCC (Meehl et al., 2007). For the HadRM3 temperature projections used in this thesis, however, such an uncertainty is not considered — all the ensemble members of HadRM3 are driven by a single emission scenario (see Section 3.4). In addition to anthropogenic emissions, there are uncertainties with natural forcing agents, including variations in solar radiation and aerosols from volcanic eruptions. Since both of these factors cannot be predicted to a useable accuracy, their effects have not been incorporated in most studies on climate prediction uncertainties (Collins, 2007; Chapter 2 in Murphy et al., 2009).

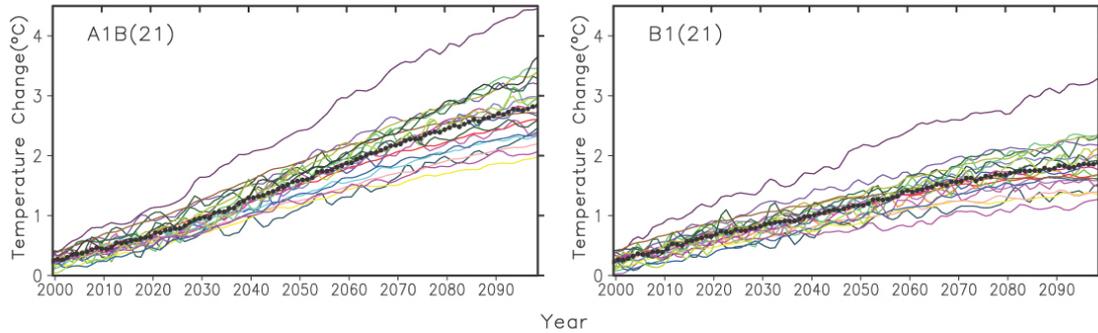


Figure 2.2: Smoothed (based on a 3-year moving average) time series of globally averaged surface air temperature change projected by 21 GCMs driven by SRES A1B (left) and B1 (right) scenarios (Nakićenović and Swart, 2000) for the 21st century relative to the average from 1980 to 1999. The black dots show the annual ensemble mean values, while the coloured lines show the results from individual GCMs. Reproduced from Fig. 10.5 of Meehl et al. (2007).

The second source of uncertainty arises from the imperfect representation of the climate system by models (model uncertainty). Some processes in the climate models are parameterized, where the parameters are determined by a mixture of theory, empirical fits in observational studies and model simulations (Collins, 2007). The uncertainty due to the choice of model parameters is known as ‘parameter uncertainty’. Additionally, there is ‘structural uncertainty’ on the choice of grid, resolution and set of processes to be included in the model (Murphy et al., 2007). These uncertainties lead to different projections for climate change under an identical emission scenario. This is demonstrated in Fig. 2.2 which shows the globally averaged surface warming projected by different models used in AR4 of IPCC (Meehl et al., 2007) for 2 SRES scenarios. For both scenarios, the spread in the warming projected by various models is comparable to the ensemble mean warming.

For regional climate projections using a RCM nested within a GCM, any errors in the driving GCM fields are carried over to the RCM, leading to potential systematic biases in the RCM simulations (Noguer et al., 1998). In addition, the formulation of RCM, including its parameters and structure, brings another tier of model uncertainty. The effects

of this uncertainty were investigated in the EU project PRUDENCE² where 10 RCMs for Europe with similar resolution (about 50 km) but different formulation were driven by the same boundary conditions provided by a single GCM (Christensen and Christensen, 2007). For example, Beniston et al. (2007) examined the projections of extreme weather events by these models and found that across different RCMs, there is considerable variation in both the signs and magnitudes of changes in summer extreme precipitation by the end of this century.

The third source of uncertainty arises from the internal variability of the climate system. This refers to natural variability due to the interactions among the atmosphere, ocean and land in the absence of anthropogenic influences. Variability on a timescale of years to decades is particularly important for climate predictions as their effects will be superimposed on the long-term climate changes by human activities (Chapter 2 in Murphy et al., 2009). El Niño-Southern Oscillation and North Atlantic Oscillation are examples of this type of variability.

The relative contributions of the above sources of uncertainties in climate predictions have been addressed in a number of studies. Rowell (2006) compared the uncertainties from various sources on temperature and precipitation projections for UK and Ireland by the end of the century, by considering results from different groups of PRUDENCE RCM simulations. For example, the uncertainty related to emissions was estimated by forcing the same group of RCMs with 2 different SRES scenarios, while the RCM formulation uncertainty was estimated by forcing different RCMs with the same boundary conditions, as described above. The projection uncertainty from the formulation of GCM was found to be the largest in all cases. Déqué et al. (2007) considered the same PRUDENCE experiments but for the whole of Europe. The uncertainty related to the choice of GCM was also found to be larger than the uncertainties from other sources, except for summer precipitation where the uncertainty from the RCM formulation becomes comparable to the choice of GCM. Hawkins and Sutton (2009) estimated the variation of fractional uncertainty (the ratio of prediction uncertainty to the predicted mean change) from each source with pre-

²Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks and Effects

diction lead time for surface decadal mean temperature projections. This was done by considering projections from 15 GCMs used in IPCC AR4, driven by 3 different SRES scenarios. For the British Isles, internal variability has the largest fractional uncertainty at shorter lead times, but scenario uncertainty gradually becomes dominant as lead time increases. Model uncertainty is comparable to scenario uncertainty by the end of the century. The results from the above studies suggest that the relative importance of different sources of uncertainties may vary in different situations.

2.3.4 Quantification of climate prediction uncertainties

In view of the above-mentioned uncertainties, probabilistic climate predictions are more useful for policymakers regarding adaptation and mitigation actions compared to single climate projections (e.g. New et al., 2007). In order to produce probabilistic predictions, each of the above uncertainties needs to be quantified and a probability distribution for the climate needs to be specified (Rougier, 2007). Different ensemble techniques have been developed to help quantify different sources of uncertainties.

The uncertainty related to internal variability can be quantified using an ‘initial condition ensemble’ of climate models, where each member has the same formulation and is driven by the same emission scenario, but is initiated with slightly different conditions (e.g. Stainforth et al., 2005).

As for model uncertainty, two ensemble techniques have been applied. The first one is the ‘multi-model ensemble’ (MME) method which uses an ensemble of GCMs from different modelling centres. Probabilistic climate predictions can then be made by making certain assumptions, such as the differences between each model in the ensemble and the true climate are independent. The ensemble members may also be weighted according to their abilities to simulate different climate variables in the past (Tebaldi and Knutti, 2007). As an example of the MME method, Tebaldi et al. (2005) uses a Bayesian approach to estimate the probability density function of future regional temperature changes based on a 9-member GCM ensemble. The second method, ‘perturbed physics ensemble’ (PPE), is the one considered in this thesis for temperature projections of London and Budapest. A

PPE consists of multiple variants of a single climate model. The set of model parameters in each variant is perturbed by expert judgement to sample a plausible range of parameter values (Collins et al., 2006). For example, Clark et al. (2006) used a 53-member PPE of a GCM to estimate the model uncertainties in the projections of extreme heat frequency, intensity and duration over the Northern hemisphere under a doubled carbon dioxide concentration scenario. Each of these two ensemble techniques has its limitations. For the MME method, the different models are assembled on an opportunity basis, which means that different models within the ensemble are not designed to systematically sample the modelling uncertainties (Tebaldi et al., 2005). On the other hand, a PPE is unable to sample the structural uncertainty as all the ensemble members belong to the same modelling framework (Chapter 3 in Murphy et al., 2009).

Finally, scenario uncertainty is difficult to quantify, since there is no consensus on how (and in fact whether) the relative likelihood of different future emissions scenarios can be assessed (Meehl et al., 2007). As a result, in practice, separate probabilistic predictions of climate change are generally produced for multiple emission scenarios (e.g. UKCP09; Murphy et al., 2009) for users' reference.

2.4 Hot weather and human health

Before discussing the adverse health impacts of extreme heat, this section first outlines a simple physical model representing the heat transfer process between a human body and the environment. This helps identify the key environmental variables causing health impacts.

2.4.1 Human response to heat

The core temperature of the human body at rest is normally about 37 °C. The thermoregulatory system of the body maintains this temperature by keeping a balance of heat transfer between the body and its surrounding environment. This balance can be represented by

the equation

$$M - W = C_n + K + R_n + E_{sw} + E_{res} \quad (2.1)$$

where each of the terms represent the rate of heat production or loss via different pathways: M is the metabolic heat production of the body, W is the mechanical work done by the body, C_n , K and R_n represent the heat loss from the body through convection, conduction and radiation respectively, and E_{sw} and E_{res} represent the heat loss by evaporation of sweat on the skin and respiration respectively (Fanger, 1970). A schematic diagram for this balance is shown in Fig. 2.3.

This image has been removed by the author of this thesis for copyright reasons.

Figure 2.3: Schematic representation of the heat balance model between a human body and the environment. Reproduced from Fig. 1 of Havenith (1999) with permission.

The efficiency for the body to lose heat in hot environments depends on a number of environmental variables, including air temperature, humidity and wind speed. The rates of heat loss through convection (C_n), radiation (R_n) and respiration (E_{res}) all decrease with increasing air temperature. When the air temperature is above the skin temperature (normally about 33 °C), heat is gained by the body through convection (Havenith, 1999). Convective heat transfer increases with wind speed. The evaporation of sweat, which is the main mechanism for the body to lose heat (Koppe et al., 2004), is affected by both humidity and wind speed. Its rate (E_{sw}) is linearly proportional to the difference of the water vapour pressure between the skin and the environment. It should be noted that another commonly used humidity variable, relative humidity, is less relevant in determining the evaporation rate. This is because relative humidity, defined as the ratio of actual vapour pressure to

the saturated vapour pressure at the observed air temperature, is a measure dependent on air temperature. Evaporation of sweat is still possible at 100% relative humidity if the ambient vapour pressure is lower than the vapour pressure on the skin (Havenith, 2005). A more quantitative discussion on how these environmental variables and clothing affect each of the heat transfer components in (2.1) can be found in Chapter 1 of Parsons (2003).

2.4.2 Heat disorders and heat-related mortality

Under heat stress, if the thermoregulatory system cannot meet the demand to lose heat in order to maintain the heat balance (2.1), acute heat illnesses may develop. These range from minor heat fatigue to the most serious form, heat stroke which has a high risk of permanent disability or even mortality as the kidney, liver and the central nervous system fail to function properly (Chapter 10 in Parsons, 2003). In addition, pre-existing cardiovascular and respiratory diseases may be aggravated in hot conditions as the human body gives priority to the thermoregulatory system. As such, among the deaths that occurred during heatwaves, only a small proportion of cases report acute heat illness as the cause of death, while most of the cases are attributed to cardiovascular and respiratory diseases (Basu and Samet, 2002; Kovats and Hajat, 2008).

Several groups of people are considered to be more vulnerable to heat. The elderly is the largest defined group at risk (World Health Organization, 2008) because of a number of reasons. The capacity of their cardiovascular system is reduced, leading to a reduction in the efficiency of heat transport from the body to the skin. In addition, the thermoregulatory system of elderly people respond to changing environments more slowly, so they are less able to sweat (Havenith, 2005). As a result, at high temperatures, mortality has been observed to increase with age in a number of studies (e.g. Hajat et al., 2007). Other susceptible groups include people who are less fit, dehydrated, overweight, with chronic diseases and taking medication. Details on the physiological aspects are discussed in Havenith (2005) and Koppe et al. (2004).

2.4.3 Adaptation to heat

Humans gradually adapt to, or *acclimatize* to, changes in their surrounding thermal environments in two different ways. When exposed to hot environments, ‘physiological acclimatization’ involves adjustments in the responses of the thermoregulatory system, such as changes in heart rate, sweat rate, sweat content and the ambient temperature at which sweats start to be produced (Chapter 10 in Parsons, 2003). The time required for these processes to complete is on the scale of days to weeks (Koppe et al., 2004). ‘Behavioral acclimatization’ includes a wide range of actions taken by individuals and the society to adapt to the thermal environment. For example, individuals adjust their clothing and activity levels during the course of the year (Chapter 2 in Parsons, 2003), while at the society level, buildings are designed to adapt to the local climate (e.g. Hacker et al., 2005).

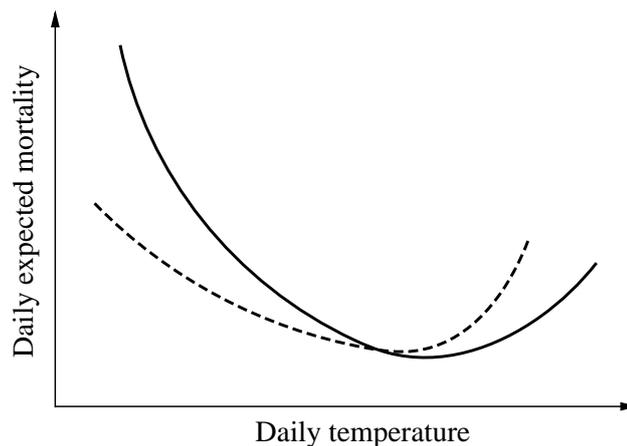


Figure 2.4: Typical relationships between daily expected mortality and air temperature at locations with warmer climates (solid line) and locations with cooler climates (dashed line).

The ability of humans to adapt to different climates leads to geographical variation in the relationships between daily expected mortality and air temperature. In subtropical and temperate regions, such relationships can be described by U-shaped or a J-shaped curves, similar to the ones shown in Fig. 2.4. Mortality is the lowest at a range of intermediate temperatures, and increases towards the two extreme ends. The slopes and turning points of

these curves vary geographically. In a comparison of mortality-temperature relationships in 11 cities in eastern United States, Curriero et al. (2002) found that expected heat-related mortality increases more rapidly with temperature in cities at higher latitudes with cooler climates, while the temperature at which minimum mortality occurs is lower in these cities. These suggest that the population living in cooler climates is generally less well adapted to high temperatures. Similar patterns of mortality-temperature relationships were also observed in Keatinge et al. (2000) which considered 7 European regions and McMichael et al. (2008) which considered 12 cities from 5 different continents.

Mortality is observed to be higher in heatwaves occurring in the earlier part of the season (Kalkstein and Davis, 1989). This is possibly related to short-term physiological acclimatization, such that people can gradually cope with higher temperatures as the summer progresses. Another possible explanation to this phenomenon is that a large part of a finite pool of susceptible people dies in the first heatwave of the season, leading to decreased mortality in subsequent heatwaves (Hajat et al., 2002). In a study of mortality related to respiratory and cardiovascular diseases in 12 US cities, Braga et al. (2002) observed that after a heat event, the number of deaths dropped below the expected level for a few days, suggesting that such a short-term ‘harvesting’ or ‘forward displacement’ in mortality has taken place. Grize et al. (2005), on the other hand, observed that the number of deaths in Switzerland remained elevated after the end of the heatwave in 2003, and concluded that ‘harvesting’ cannot explain all of the increased mortality in hot spells.

2.5 Previous assessments of climate change impacts on heat-related mortality

There have been several published studies assessing the impacts of climate change on heat-related mortality. As mentioned in Section 1.1, while these studies have adopted a similar general approach involving health impact modelling and climate modelling, there are differences in certain aspects of their models and treatment of uncertainties. This section reviews four of these studies, with a focus on such differences. Although some

other mortality projections (e.g. Guest et al., 1999; Davis et al., 2004; Hayhoe et al., 2004 and Knowlton et al., 2007) are not included in this review for the sake of conciseness, the four studies discussed below are representative of the various methodologies adopted in this area of research.

Kalkstein and Greene (1997) projected the excess heat-related mortalities in 44 US cities for two decades centred around 2020 and 2050 respectively. A ‘synoptic classification’ scheme, which classifies days into different air mass categories using the discriminant analysis technique based on past meteorological observation data (Sheridan, 2002; Chapter 13 in Wilks, 2006), is central in this work. For each city, the air mass categories causing increased mortalities were first identified. For days belonging to each of these ‘high risk’ air mass types, multiple regression models were then fitted to the daily excess mortality and meteorological data. The explanatory variables include measures of temperature, humidity, wind speed, cloud cover, etc., and the choice of these variables is different for each location. Climate projections from 3 different GCMs were used. In order to account for the possible effect of acclimatization by the population, an ‘analogue city’ approach was used in the projection of mortalities. For each city, based on the future climate projected by a GCM, an ‘analogue city’ which has a similar climate at the present is identified. An example given by Kalkstein and Greene (1997) is that if the future GCM-projected climate for New York City is similar to the present-day climate of St. Louis (a city located 1400 km to the southwest of New York City), St. Louis is the ‘analogue city’ for New York City. The future excess heat-related mortalities under each ‘high risk’ air mass types at each city are projected using the mortality model developed for the corresponding air mass types in its analogue city. For the above example, future mortalities in New York City are projected using the mortality model for St. Louis. This approach assumed that the future response to temperatures by the population in a certain city would be the same as how the population at the corresponding ‘analogue city’ respond at the present. The total excess heat-related mortalities for each summer totalled for all 44 cities were projected to increase by 73 to 158% (depending on the choice of GCMs) from 1990s to 2050s. There was no mention by Kalkstein and Greene (1997) that the possible biases in the GCM projections had been accounted for. In addition, it is difficult to assess the

validity of this ‘analogue city’ assumption. Given that the relationship between mortality and weather is location-specific, the mortality projections are likely to be sensitive to the choice of ‘analogue city’.

Donaldson et al. (2002) projected heat-related mortalities for the UK as part of a report by the Department of Health on the health effects of climate change in the country. This report is based on the UKCIP98 climate projections³ for the UK using a GCM. Donaldson et al. (2002) developed a single mortality regression model using the daily heat-related mortality in the whole of UK and daily Central England Temperature (CET)⁴. Details of the specification of the regression model were not provided. The projected changes in mean summer temperature in the UK under 4 emission scenarios for 3 future periods (2020s, 2050s and 2080s) were added to currently observed values of CET to obtain the future CET series (a ‘change factor’ approach; see Chapter 5 for details). Future heat-related mortalities were then projected using the mortality model driven by the future CET series. Assuming no change in population and no adaptation, the annual mean number of heat-related deaths in the UK was projected to rise from 798 in the 1990s to 2793 (an increase of 250%) in the 2050s under the ‘medium-high’ emissions scenario. For the other 3 scenarios, the projected number of deaths ranges from 1368 (an increase of 71%) to 3249 (an increase of 307%). These are rather crude projections given that the mortality-temperature relationship for the whole of UK is described by a single mortality model which considers only the CET. It is unlikely that the local effects of extreme high temperatures of mortality can be well captured by the model. In addition, as will be demonstrated in Chapter 5, it is not justifiable to project future temperatures by only considering a shift in the mean of the current temperature distribution, as was done by the authors. This is because other properties of the temperature distribution, such as variance, can also change with time.

Heat-related mortalities in Lisbon, Portugal for 2020s, 2050s and 2080s were projected by Dessai (2002, 2003). Daily observed mortality and maximum temperature data were

³This refers to the report of UK Climate Impacts Programme published in 1998. UK Climate Projections UKCP09 described above are the projections for UKCIP published in 2009 (Murphy et al., 2009).

⁴CET is the weighted mean temperature observed in multiple sites in the UK and is representative for a triangular area enclosed by Lancashire, London and Bristol.

These images have been removed by the author of this thesis for copyright reasons.

Figure 2.5: Mortality model fits for (a) Lisbon by Dessai (2002) and (b) London by Gosling et al. (2007). In (a) the model fits shown are for two decades as labelled. In (b) the model fit is shown by the blue line, while the red lines are the 95% bootstrap confidence interval limits for model predictions. Reproduced from Fig. 4 of Dessai (2002) and Fig. 2 of Gosling et al. (2007) respectively with permission.

used to construct the mortality model. In the model, the daily numbers of excess deaths were first grouped into 1 °C intervals according to the maximum temperature observed on the day mortality occurs. The average daily number of excess deaths for each interval was then calculated. This became the response of the non-linear mortality model, while the explanatory variable was the lower bound temperature value of each interval. Examples of model fits are shown in Fig. 2.5(a). Two RCMs with horizontal resolution of about 50 km forced by the same GCM were used in the study. Similar to Donaldson et al. (2002), future temperature projections were obtained by considering a change factor in the mean, but here the change factor varied from day to day based on an interpolation of the monthly temperature anomalies projected by each RCM. The annual mean summer heat-related mortality rate (per 100,000 population) was projected to increase from 5.4 in the period 1980 to 1998, to 16.2 and 29.5 (increases of 200% and 450%) in 2050s, for the two RCMs respectively. Dessai (2003) considered the sensitivity of mortality projections to effects of adaptation by repeating the projections, but assuming that the mortality-temperature function estimated by the mortality model will shift by 1 °C. For example, the predicted mortality for a day in the future with maximum temperature of 30 °C will be the value

predicted by the regression model at 29 °C (see Section 6.2.2 for further discussion). With this assumption, the projected mortality rates for 2050s became 7.3 and 13.4 (increases of 35% and 148%) for the two RCMs respectively. These results suggest that both the choice of RCMs and the incorporation of adaptation effects have significant impacts on the mortality projections. The use of average number of deaths in temperature intervals as the response in the mortality model effectively smoothed out the variance in daily mortality counts. It is then difficult to assess the uncertainties in the daily predictions of the mortality model. In addition, with fewer mortality observations at extremely high temperatures, the model prediction may become less precise.

The mortality projections by Gosling et al. (2007, 2009b) used the GCM HadCM3 driven by SRES A2 scenario and covered six cities: London, Budapest, Lisbon, Boston (US), Dallas (US) and Sydney (Australia). Separate mortality models were fitted for each city. The specification of the models were similar to the model in Dessai (2002), but with a temperature interval width of 2 °C [see Fig. 2.5(b) for the model fit for London]. Gosling et al. (2009b) adopted a different method to calibrate future HadCM3 projected maximum temperature for each location. Logistic distributions were fitted to observed and GCM simulated temperatures for the present-day and GCM projected temperatures for the future⁵. The (unknown) future observed temperatures are also assumed to follow a logistic distribution. The location and scale parameters of the future observed temperature distribution are estimated by adding the changes in the respective parameters for the GCM temperatures between the present-day and the future to the parameters for present-day observed temperatures. We will revisit this calibration method in Section 5.2.4 as the temperature calibration framework for this thesis is discussed. Using these estimated location and scale parameters, a random observed temperature series for the future was produced and was used to drive the mortality models. For London, the annual mean summer heat-related mortality rate per 100,000 population was projected to increase from 1.8 per 100,000 population during 1961 to 1990, to 12.5 per 100,000 population (an increase of 590%) during 2070 to 2099, while the mortality rate for Budapest increases from 5.4

⁵Logistic distributions are symmetric distributions but have heavier tails compared to normal distributions. See Chapter 5 in Davison (2003).

to 98.5 per 100,000 population (an increase of 1700%). Adopting the same approach as in Dessai (2003), the effects of adaptation on mortality projections were assessed. In addition, the sensitivity of mortality projections to the use of projected temperatures based on an alternative emission scenario (SRES B2) was also investigated.

Several aspects of the approach adopted in this thesis are different from the studies reviewed above. First, temperature projections are from HadRM3 which has a finer horizontal resolution (with minimum grid spacing of about 25 km). This should allow a better representation of local variation of temperatures and the extremes. Second and most importantly, in order to account for the biases in the RCM simulations, more transparent methods of calibrating RCM projected temperatures will be considered (Chapter 5). Third, for the uncertainty analysis of mortality projections (Chapter 6), although scenario uncertainty cannot be considered, this thesis will assess the sensitivity of extreme mortality projections to climate model uncertainty, using the PPE of HadRM3. Previous studies have not addressed this uncertainty, which as discussed in Section 2.3.3, is potentially a major source of uncertainty in climate projections.

2.6 Summary

The review of background literature in this chapter has covered various areas, including climate and its predictions, health and previous heat-related mortality projections. Some of the important points which are especially relevant to the work presented in the following chapters are:

- In addition to the large-scale circulation patterns, soil moisture has an important role in the occurrences of heatwaves in Europe;
- Uncertainties of climate predictions arise from three sources: uncertainty in future emissions, climate model uncertainty and internal variability of climate. The perturbed physics ensemble method is used to estimate the climate model parameter uncertainty;
- Regional climate models provide better spatial representation of climate features,

but like global climate models, there will be systematic biases in their simulations. The formulation of the regional climate model is another source of climate model uncertainty;

- Humans adapt to different climates through changes in physiological and behavioral responses.

Chapter 3

Data

3.1 Aim

This chapter provides an overview of various data sets used in this thesis. These include the demographic and meteorological station data for developing the statistical heat-related mortality models for the two focus cities, London and Budapest (Chapter 4), the gridded European temperature observation data set for calibrating projections from the regional climate model HadRM3 (Chapter 5) and different simulations of HadRM3 to be considered for projecting changes in extreme daily temperatures and mortalities (Chapters 5 and 6).

3.2 Two focus cities: London and Budapest

3.2.1 Meteorological observation station data

Sub-daily meteorological station data to be used are observed at London Weather Centre (51.521° N, 0.11° W, 43 m above sea level) in central London and Lőrinc (47.433° N, 19.183° E, 138 m above sea level) in the suburban area of Budapest from 1991 to 2005 (UK Met Office, 2006). The measurements are made on a rooftop for London Weather Centre and near the ground for Lőrinc (Jones and Lister, 2009). Observed variables in-

clude air temperature (in °C), vapour pressure¹ (in hPa) and wind speed (in m s⁻¹). The choice of using vapour pressure as the humidity variable is justified in Section 2.4.1. Observations are available at an hourly interval for London and at a 3-hourly interval for Budapest. There are about 0.25% of observations missing in each data set. In this thesis, these missing observations are filled in by linear interpolation in time. However, if there are four or more consecutive missing hourly observations (London) or two or more consecutive missing 3-hourly observations (Budapest), the whole day is omitted in the analysis. This is because the diurnal variation in weather variables may be missed out if interpolation was done. Daily mean values of air temperature T^s , vapour pressure e^s for both cities are then calculated by averaging the hourly or 3-hourly values.

Time series of T^s and e^s for the two stations are shown in Figs 3.1 and 3.2 respectively. Only the periods where mortality data (described in Section 3.2.2) are available are shown for easier comparison. There are no apparent trends in the temperature series. For the London temperature series [Fig. 3.1(a)], the sharpest summer peak occurs during the 2003 heatwave, when the daily mean air temperature reached 29.3 °C on 10th August. The annual cycles in daily mean vapour pressure [Figs 3.1(b) and 3.2(b)] for both stations have strong associations with the corresponding cycles in the air temperature series.

3.2.2 Demographic data

Data of daily elderly (age 65 or above) mortality counts from all causes (denoted by Y) for the London government office region (Greater London) from 1993 to 2003 are obtained from the UK Office of National Statistics, while such data for Budapest from 1992 to 2001 are obtained from the European Union's 'Assessment and prevention of acute health effects of weather conditions in Europe' (PHEWE) project (Michelozzi et al., 2007). The time series are shown in Figs 3.1(c) and 3.2(c) respectively. At both locations, elderly mortalities appears to have slow overall decreasing trends, and their annual cycles are generally negatively associated with the daily mean air temperature annual cycles [Figs 3.1(a) and 3.2(a)]. The mortality annual cycles in Budapest appear to be weaker than that

¹Vapour pressure is derived from measurements using dry bulb and wet bulb thermometers.

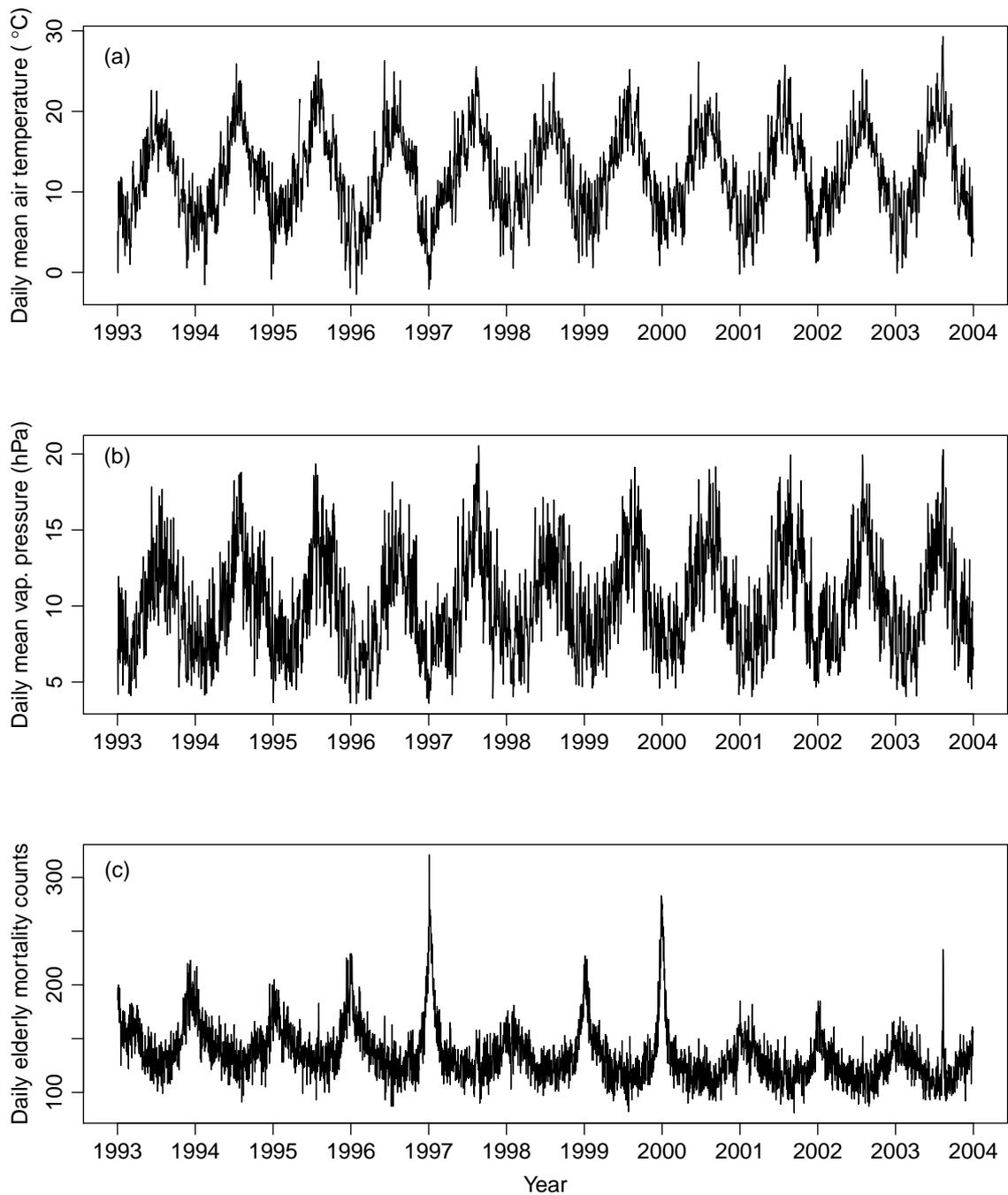


Figure 3.1: Time series of observed (a) daily mean air temperature T^s ; (b) daily mean vapour pressure e^s ; (c) daily elderly mortality counts in London Y from 1993 to 2003. Observations of T^s and e^s are made at London Weather Centre station.

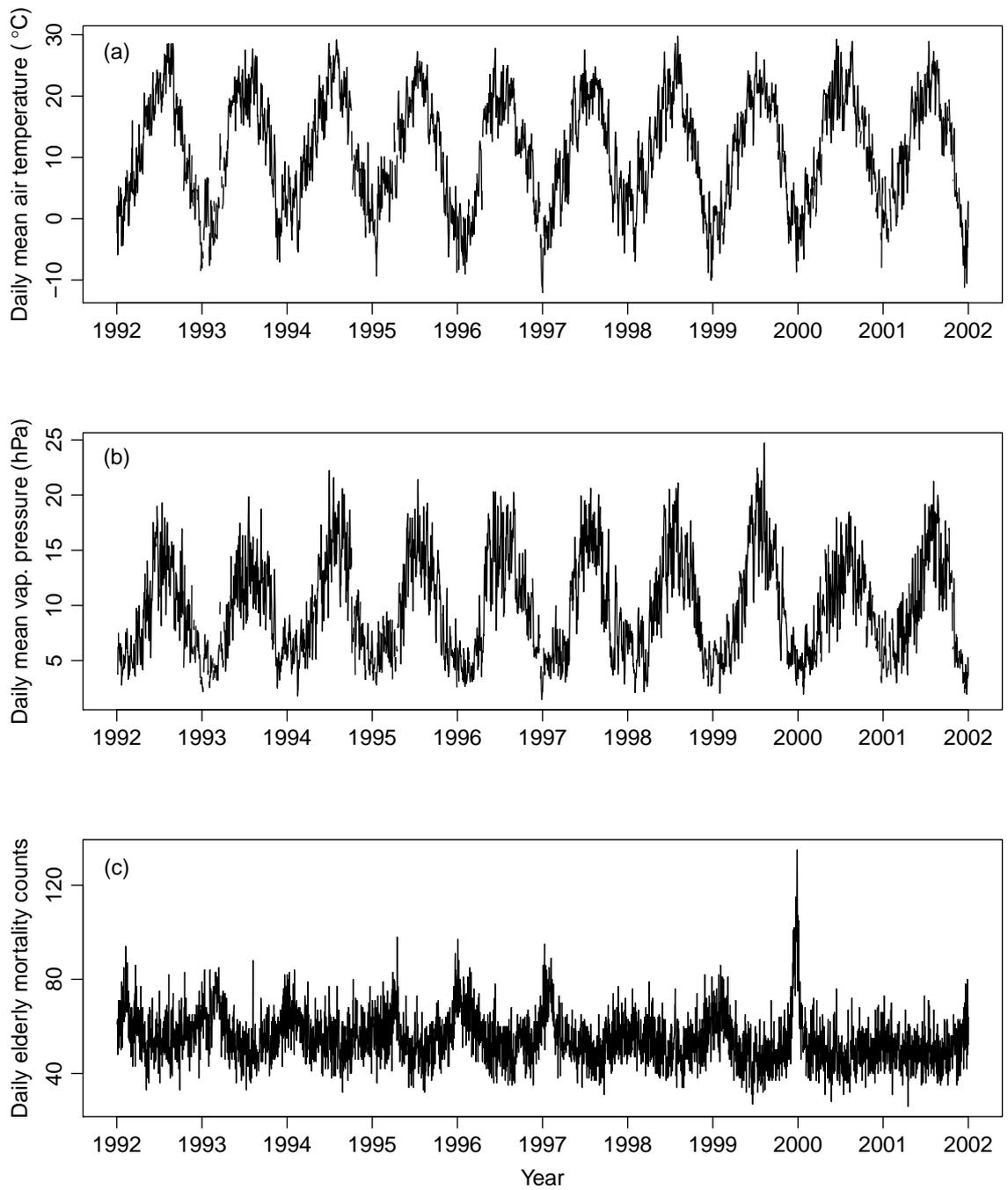


Figure 3.2: Same as Fig. 3.1 but for Budapest from 1992 to 2001. Observations T^s and e^s are made at Lőrinc station.

in London, especially towards the latter part of the series, even though there is no apparent change in the amplitude of daily mean air temperature annual cycles in Budapest throughout the time series [Fig. 3.2(a)]. Two distinct summer mortality peaks can be observed in London, one in 1995 and the another in 2003.

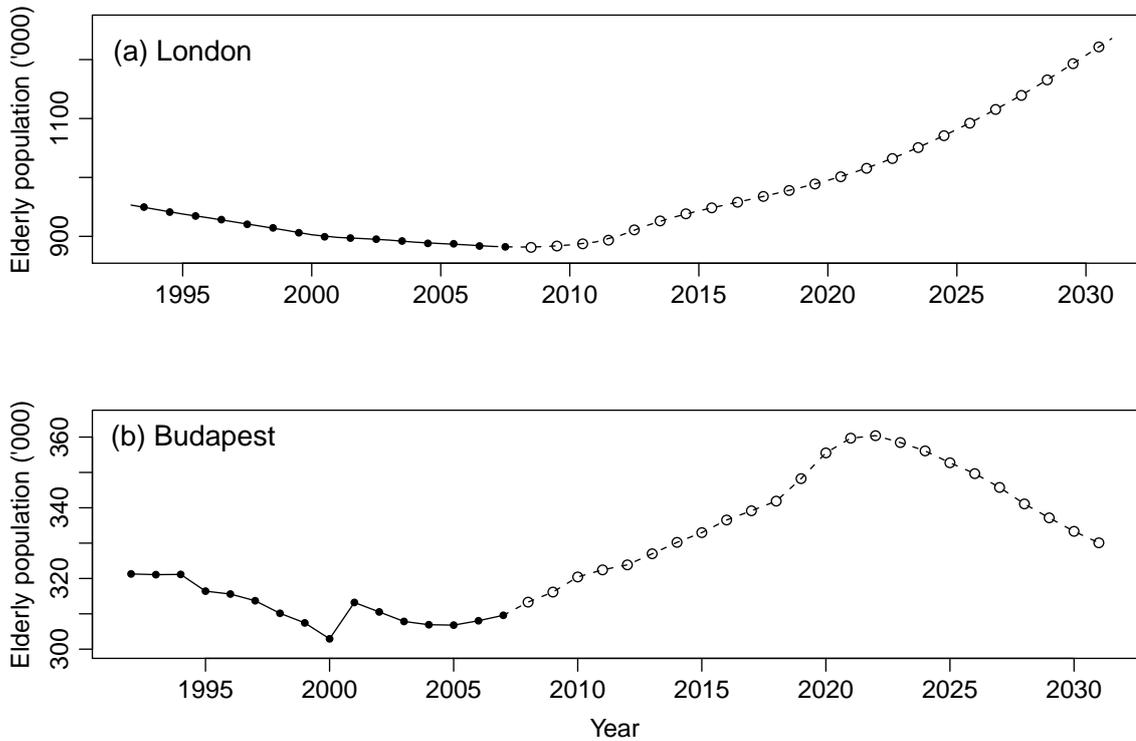


Figure 3.3: Time series of annual estimates (shown by circles) of past and future projected elderly population (in thousands) in (a) London and (b) Budapest. Linear interpolation of these annual estimates in time are shown by solid (past) and dashed (future) lines.

In order to take account for demographic changes when fitting the statistical mortality models, annual elderly population estimates for each location from 1992 to 2007 are obtained (UK Office of National Statistics, 2008b; Hungarian Central Statistical Office, 2009). These are shown by filled circles in Fig. 3.3. For the periods where mortality data are available, apart from a jump in Budapest between 2000 and 2001, the sizes of elderly population at both locations were slowly decreasing despite the general ageing trend in Europe (Chapter 1 in Eurostat, 2008). A possible reason is that migration of the elderly out of the city areas dominates the population changes (Chapter 1 in UK Office of Na-

tional Statistics, 2007). However, in more recent years, the decrease in elderly population sizes at both locations has slowed down. For the purposes of fitting heat-related mortality models, the annual population estimates are interpolated linearly to give daily estimates, denoted by P and are shown by solid lines in Fig. 3.3.

The projected changes in future elderly population will be considered when projecting the changes in heat-related mortality (Section 6.4). For London, the 2006-based subnational UK population projection up to 2031 will be used (UK Office of National Statistics, 2008c). For Budapest, since no population projection data for the city are publicly available, these are derived from the 2004-based subnational population projection data for larger regions within Hungary up to 2031 (Eurostat, 2009). The details of the derivation are given in Appendix A. As shown by the dashed lines in Fig. 3.3, the size of elderly population in London is projected to increase gradually in the next two decades, while for Budapest, the elderly population size will increase before falling again in the early 2020s. These population projections are subject to uncertainties because they are produced using assumptions on mortality rates, internal (within-country) and international migration based on current trends (UK Office of National Statistics, 2008a).

3.3 European gridded temperature observations E-OBS

The observed daily mean screen-level air temperatures from the European daily gridded land surface observational data set E-OBS Version 2.0 (denoted by T^o ; Haylock et al., 2008) produced by the EU-FP6 project ENSEMBLES (van der Linden and Mitchell, 2009), are used in the evaluation and calibration of RCM temperature projections. This data set covers the period from 1950 to 2008. As an example, the observed mean air temperature on 1 August 1995 is shown in Fig. 3.4(a). The main advantage of using this data set is that its observations can be directly compared to the output from HadRM3 because it is available in a grid identical to that used in HadRM3 - with a rotated north pole at 39.25° N, 162° W and a spatial resolution of $0.22^\circ \times 0.22^\circ$, i.e. about 25 km on the equator of the rotated grid². In addition, to be consistent with the RCM output, the gridded

²The use of a rotated grid in RCMs allows a more uniform horizontal resolution in the model domain.

daily temperatures represent grid box averages rather than point observations.

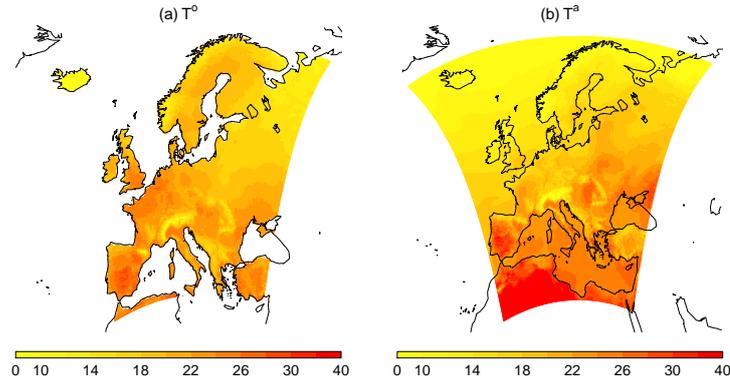


Figure 3.4: Mean air temperature (in °C) on 1 August 1995 from (a) E-OBS data set (T^o) and (b) simulations of HadRM3 forced by ERA-40 (T^a). Values of T^a for grid boxes over waters are also displayed here to show the entire spatial domain of HadRM3.

Gridded daily mean temperatures for the E-OBS data set were produced by first interpolating the daily mean temperatures observed from about 1200 stations across Europe to a 0.1° master grid using the kriging method (Chapter 3 in Cressie, 1993), then averaging to give grid box averages on the 0.22° grid with about 18,000 boxes. Details on the density of the observation stations and interpolation methods are described in Klok and Klein Tank (2009) and Haylock et al. (2008) respectively. Hofstra et al. (2009) shows that over areas with fewer observation stations, for example Spain, Turkey, Sweden and western Russia, temperature observations are over-smoothed in the interpolating process, resulting in an under-estimation of the grid box average daily values, especially for the extremes. This should be noted when extreme T^o and HadRM3 simulations over such regions are compared in Chapter 5.

3.4 Regional climate model HadRM3

The RCM used in this thesis, HadRM3, has a spatial domain covering UK and Europe [see Fig. 3.4(b)], with 170 grid boxes on the zonal direction and 190 grid boxes on the meridional direction. As mentioned in Section 3.3, the minimum grid box length is about 25 km. With this spatial resolution, the area of Greater London is resolved into more than

one grid box [Fig. 3.5(a)]. This RCM is nested within the Hadley Centre coupled ocean-atmosphere GCM HadCM3 (Gordon et al., 2000), which has a horizontal resolution of $2.5^\circ \times 3.75^\circ$ (about 300 km at the mid-latitudes) with 19 vertical levels for the atmosphere and $1.25^\circ \times 1.25^\circ$ with 20 vertical levels for the ocean. Refer to Fig. 2.1 for an illustration of the different resolutions between HadRM3 and HadCM3. HadRM3 is based on an improved version of the atmospheric component of HadCM3 (Jones et al., 2004; Pope et al., 2006). As it does not have an interactive ocean component, sea surface temperatures and sea ice extent are prescribed from HadCM3. Details of the nesting technique and model formulation of HadRM3 are described in Jones et al. (1995) and Buonomo et al. (2007) respectively. It should be noted that HadRM3 does not include any influences of the urban surfaces and built environment on the climate (Annex 7 in Murphy et al., 2009), therefore the ‘urban heat island’ effect (Wilby, 2007) cannot be represented.

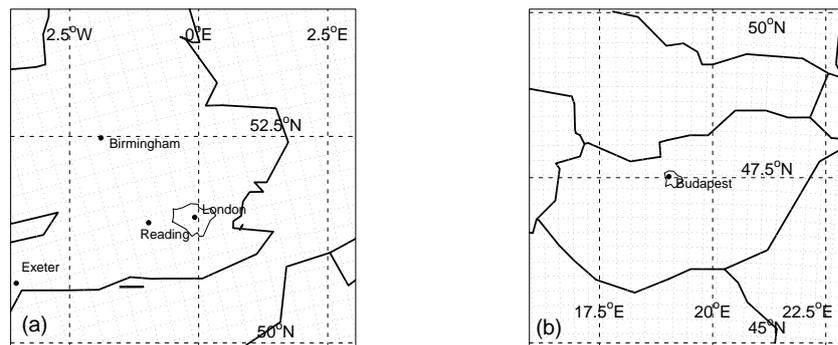


Figure 3.5: Maps illustrating the spatial resolution of HadRM3 over (a) UK and (b) Hungary. The grey dotted lines are boundaries of the grid boxes. The thin black lines are the boundaries of Greater London and Budapest, with the dots within showing the locations of London Weather Centre and Lőrinc.

3.4.1 Description of model simulations

In the main simulation, the parameter settings of HadRM3 are consistent with the standard version of HadCM3 described in Gordon et al. (2000). HadRM3 is driven at the bound-

aries of the domain by 6-hourly values of surface atmospheric pressure, wind, temperature and humidity from the output of HadCM3. Future greenhouse gases concentration from the IPCC SRES A1B scenario is used to drive HadCM3. This scenario assumes a world with rapid economic growth, with population peaking by 2050 and with a balanced use between fossil and non-fossil energy sources (Nakićenović and Swart, 2000). This HadRM3 simulation runs on a twelve 30-day months calendar from 1950 to 2099. To match the definition of summer (15th May to 15th September) on the Julian calendar stated in Section 1.2, summer for T^g is defined to be the 135th to 254th day of each model year. Daily mean air temperatures from this simulation, which are denoted by T^g , will be used in the calibration of extreme temperature projections over Europe.

In the second simulation, HadRM3 is driven at the boundaries by the ERA-40 re-analysis from the European Centre for Medium-Range Weather Forecasts (ECMWF) (Upala et al., 2005). This simulation runs from December 1957 to November 2001. If the re-analysis data are considered to be ‘perfect’ boundary forcings, a comparison of simulated daily mean temperatures from this run, denoted by T^a , with the E-OBS observations T^o , can help identify the model biases related to the RCM formulation. The T^a output on 1 August 1995 is shown in Fig. 3.4(b) as an example. Compared to T^o on the same day [Fig. 3.4(a)], while the lower temperatures in elevated areas are well simulated by the RCM, T^a is about 2 to 3 °C lower over UK and Scandinavia. A detailed comparison of the T^o , T^a and T^g distributions will be presented in Chapter 5.

3.4.2 Perturbed physics ensemble

In the assessment of the effects of climate model uncertainties on extreme temperature and mortality projections (Sections 5.4.3, 5.5.3 and 6.4.1), an 11-member perturbed physics ensemble (PPE) of HadRM3 simulations of T^g will be used. The same ensemble has been used to downscale GCM outputs for UK Climate Projections (UKCP09; Murphy et al., 2009). The first member is the ‘standard’ HadRM3 simulation described above. For each of the other 10 variants, atmosphere, sea-ice and land surface parameters in the driving HadCM3 model are perturbed from the standard value, while some options within their pa-

parameterisation schemes are switched on or off. The HadRM3 PPE member is then driven at the boundaries by the output of the corresponding HadCM3 simulation. The parameter settings of each HadRM3 member are also consistent with that of the corresponding HadCM3 simulations (Chapter 3 in Murphy et al., 2009). The choices of parameter perturbation for these 10 members are based on results of previous experiments with larger PPEs (e.g. Murphy et al., 2004; Collins et al., 2006), with the aim of simulating physically plausible climate variability while spanning a wide range of possible parameter values. As a reference, the values of climate sensitivity (equilibrium change in global mean temperature for doubled CO₂ concentration) of the slab models (HadSM3)³ which have the same parameter perturbations as the 10 HadCM3 variants range from 2.58 K to 7.11 K (see Fig. 3.6). The same historical and future greenhouse gases concentration (SRES A1B scenario) used in the standard simulation is also used to drive the simulations of the PPE variants. It is important to note that the known climate model uncertainties are not fully sampled in this PPE, including structural errors in the driving model HadCM3 which could be remedied by using an alternative climate model and uncertainties from the carbon cycle, sulphur cycle and ocean transport processes (Chapter 5 in Murphy et al., 2009).

The boxplots shown in Fig. 3.6 can be used to visually compare the distributions of summer T^g simulated by the ensemble members for the grid boxes covering London and Budapest. For London, while the different widths of the boxes indicate differences in the scale of T^g distributions among the ensemble members, all of the members simulate positively skewed T^g distributions, as shown by the large number of T^g values lying outside 1.5 times the interquartile range (IQR) from the upper quartile. However, comparing with the boxplot of T^o for the same period, this skewness is not seen for the distribution of T^o . For Budapest, the T^g distributions for all of the ensemble members are more symmetric, but the range of the median of simulated T^g values is greater compared to London. In Chapter 5, the calibration of future temperature projections by the PPE will be based on the comparison of T^o and each ensemble member of T^g .

³HadSM3 has a simpler configuration of the ocean compared to HadCM3.

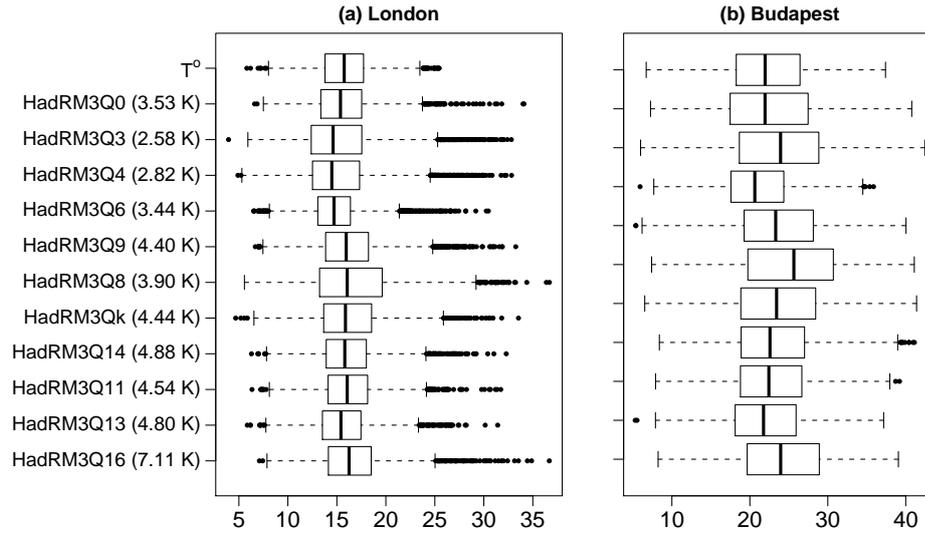


Figure 3.6: Boxplots of summer T^o and T^g (in $^{\circ}\text{C}$) simulated by HadRM3 perturbed physics ensemble members from 1970 to 1999 for the grid boxes covering the (a) London and (b) Budapest observation stations (see Fig. 3.5). Each ensemble member is labelled by its model name. HadRM3Q0 is the standard version while the others are the ensemble variants. Values of the climate sensitivity of the corresponding slab models are also shown (see text for explanation). The lower and upper boundaries of the boxes represent the lower and upper quartiles of the T^o and T^g distributions respectively. The vertical lines within the boxes show the median. Values lying outside 1.5 times the interquartile range (IQR) from the lower and upper quartiles are shown by dots.

3.5 Summary

This chapter has provided background information on the data sets used in this thesis. Some of the limitations of the data sets are highlighted, for example the ability for extreme temperatures to be represented in the E-OBS gridded observations and the range of uncertainties represented by the HadRM3 ensemble. These should be taken into account when interpreting the results presented in the following chapters.

Chapter 4

Statistical modelling of heat-related mortality

4.1 Aim

This chapter presents the specification, fitting and evaluation of the statistical mortality models which will be driven by HadRM3 output for projecting future heat-related mortalities in London and Budapest. It starts with a brief review on the modelling strategies adopted by previous research and a discussion on how the models for London and Budapest should be specified. A number of criteria for evaluating the models are then proposed. This is followed by a description of the candidate models which are fitted to the data sets described in Chapter 3. The models are then compared according to the proposed criteria, in particular their performance in predicting extreme daily mortality counts. Such comparison will be used to inform choices about mortality models.

4.2 Modelling strategies

There is extensive literature on the subject of heat-related mortality modelling as the use of such models is not limited to climate change impact assessment. In epidemiological studies, regression models are generally fitted to mortality data and meteorological obser-

vations over a long period of time (typically over 5 years) at different locations, in order to study how the mortalities of different age or socio-economic groups and mortalities from different causes vary with environmental variables (e.g. Anderson and Bell, 2009; Gouveia et al., 2003; Hajat et al., 2007; Ishigami et al., 2008). In heatwave episode analyses (e.g. Fouillet et al., 2008; Grize et al., 2005; Tan et al., 2007), similar models are used to retrospectively estimate the ‘excess’ number of deaths in a specific heatwave and to study the daily changes of mortality during the event in detail. Statistical mortality models are also employed as a part of operational heat-health warning systems in more than 20 cities worldwide (Sheridan and Kalkstein, 2004; Nogueira and Paixão, 2008). Warnings are issued if the number of deaths predicted by the model on a certain day exceeds a pre-determined threshold.

In previous published research, heat-related mortality models for the above purposes have adopted various modelling approaches. The approaches mainly differ in their choices of environmental covariates (explanatory variables) and how the systematic dependence of the response on these variables are modelled. Before these two areas are discussed, a general framework of heat-related mortality models is first described.

4.2.1 Model framework

When modelling heat-related mortalities, the response variable is usually the observed daily number of deaths at a specific location Y_i , where the subscript i is used to denote the observation on each day. As this is count data, the use of normal distributions is not appropriate. A generalized linear model (GLM; Nelder and Wedderburn, 1972), which allows the response to have non-normal distribution, is therefore commonly used.

Suppose there is a set of p covariates $\mathbf{X}_i = \{X_{i1}, X_{i2}, \dots, X_{ip}\}$. The random component of a mortality GLM generally assumes $Y_i|\mathbf{X}_i$ to have a Poisson distribution. This distribution is a member of the exponential family with a mean rate of μ_i :

$$Y_i|\mathbf{X}_i \sim \text{Poisson}(\mu_i), \quad (4.1)$$

where $\mu_i = \mathbb{E}(Y_i) = \text{var}(Y_i)$, i.e. the expectation (population mean) and variance of Y_i are equal. This Poisson distribution assumption is used since for the range of \mathbf{X}_i observed,

mortality can be considered as a rare event given the small number of people who die on each day out of the total population. The mean mortality rate then depends on the covariates through a linear predictor η_i which is the systematic component of the model:

$$\eta_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} = \mathbf{X}_i \boldsymbol{\beta} \quad (4.2)$$

where β_j ($j = \{1, 2, \dots, p\}$) are unknown parameters to be estimated. The random and systematic components of the GLM are related by a link function, which is a non-linear transformation of the linear predictor:

$$\eta_i = g(\mu_i) = \mathbf{X}_i \boldsymbol{\beta}.$$

For a Poisson response, a log link function $g(\mu_i) = \log(\mu_i)$ is the natural canonical choice (Nelder and Wedderburn, 1972).

Another type of statistical model commonly used for modelling heat-related mortality is the generalized additive model (GAM; Hastie and Tibshirani, 1990). It has the same random component as a GLM, but the systematic component involves a sum of unknown smooth functions of covariates:

$$\eta_i = g(\mu_i) = \mathbf{X}_i \boldsymbol{\beta} + \sum_{k=1}^p f_k(X_{ik}) \quad (4.3)$$

where the first term of the right hand side includes any strictly parametric model component, and f represents non-parametric smooth functions of covariates. It can be seen by comparing (4.2) and (4.3) that the main difference between GLMs and GAMs is that the latter allow a more flexible (non-parametric) specification for the dependence of the response on the explanatory covariates.

Further details on GLMs and GAMs, including the estimation of parameters and smoothing functions, are described in Appendix B. We now turn our attention to how the systematic components of mortality models are specified.

4.2.2 Choice of environmental covariates

Theories of heat exchange between the human body and the environment described in Section 2.4.1 can help provide a sound physical basis for choosing the explanatory environmental covariates in mortality models. Since air temperature, humidity and wind speed

are three variables that directly affect such heat exchange, they are possible covariates to be considered.

For air temperature, some mortality modelling studies (e.g. Fouillet et al., 2007; Gosling et al., 2007) choose daily maximum or minimum air temperature as a covariate, but daily mean air temperature is a more common choice (e.g. Ballester et al., 1997; Kunst et al., 1993). Hajat et al. (2002) argues that using the daily mean value is more appropriate because it better represents the thermal exposure experienced by humans throughout the day and night.

The inclusion of humidity is less common, possibly because continuous and reliable measurements of humidity are less readily available compared to air temperature. For studies which include humidity, relative humidity is often chosen as the covariate, for example in Armstrong (2006), Braga et al. (2001), Donaldson et al. (2003) and Kovats et al. (2004). However, as discussed in Section 2.4.1, another humidity variable, observed vapour pressure, is the more relevant variable in determining the effects of humidity on human heat stress, and will be used in one of the models to be fitted in Section 4.4. A number of studies included other daily measured meteorological variables, such as wind speed (Kunst et al., 1993) and barometric pressure (O'Neill et al., 2003). There is no consensus, however, on whether these variables have statistically significant effects on daily mortality.

Measured variables of ambient air pollution, such as concentration of fine particles, ozone, sulphur dioxide and nitrogen oxide, are also considered as potential confounders in some heat-related mortality models, for example in Baccini et al. (2008), Hajat et al. (2007) and Vaneckova et al. (2008). The concentration of certain air pollutants (e.g. fine particles and ozone) tends to be higher during extended periods of hot weather (Anderson et al., 2002). The interaction between air pollution and weather variables should therefore be considered in mortality models (Roberts, 2004). However, air pollution variables will

not be included in the models for this thesis due to data availability issues¹ and because there are also no long-term projections of regional air pollution concentration available.

A number of mortality models choose daily values of biometeorological indices as a covariate instead of air temperatures. Examples include Davis et al. (2003) and O'Neill et al. (2003) which use 'apparent temperature', Smoyer-Tomic and Rainham (2001) which uses 'humidex' and Laschewski and Jendritzky (2002) which uses 'perceived temperature'. Developed from physical models of heat exchange between the human body and its surrounding environment, these indices combine various meteorological variables, including air temperature, humidity, wind speed, solar radiation, etc., to provide measures of human sensation (Chapters 8 and 10 in Parsons, 2003; Steadman, 1979a,b, 1984). It is hoped that these indices can better represent the combined effect of meteorological variables on human health. However, without evaluating against a model which only includes air temperature as a covariate, it is unclear whether using such indices gives more explanatory power.

The possible delayed or lagged effects of environmental variables on mortality are commonly considered in the model by including these variables measured on the day when mortality occurs and also on a number of preceding days, or more simply by using an multiple-day average value of the variable concerned as a covariate in the model (e.g. Baccini et al., 2008). Schwartz (2000) introduced the use of parametric 'distributed lag models' to study how the impact of exposure to air pollution on daily mortality (response) varies with time. This was extended to non-parametric GAMs by Zanobetti et al. (2000). 'Distributed lag models' were applied to model heat-related mortality by a number of studies, such as in Armstrong (2006), Braga et al. (2001) and Hajat et al. (2005).

¹Measurements of air pollution concentration in London from the UK National Air Quality archive (http://www.airquality.co.uk/data_and_statistics.php) for the period under investigation contain a rather large proportion (about 4%) of missing values. A large number of days will have to be omitted if they are used.

4.2.3 Measurement locations of environmental covariates

Ideally, the environmental variables in a mortality model should be representative of the conditions to which the population is exposed. In practice, however, the modeller has to use the available routine observations from meteorological stations, which may be situated away from where most of the population reside (Basu and Samet, 2002). In addition, temperatures in the urban areas of the city are typically higher than that in the rural areas, especially at night (Wilby, 2003). The storage of heat by buildings, reduced wind speed caused by urban structures and anthropogenic heat production by human activities contribute to this ‘urban heat island’ effect (Oke, 1987). If there is a large variation of temperatures across a city, the relationship between mortality and environmental variables estimated by the model will be sensitive to the location of chosen meteorological stations (de’Donato et al., 2008).

This issue needs to be considered in this study because the statistical mortality model will be driven by the output from HadRM3 which represents averages over a grid box, while the meteorological stations data (T_i^s and e_i^s) described in Section 3.2.1 are point observations within the cities of London and Budapest. Considering the sizes of the areas from which mortality data are recorded (Fig. 3.5), it may be more appropriate to use gridded E-OBS data T_i^o as they may better represent the overall temperatures experienced by the London and Budapest populations. On the other hand, as described in Section 3.3, extreme T_i^o may be underestimated because of oversmoothing of station observations. If T_i^o is used in the mortality model, there could be bias in the estimated relationship between mortality and extreme temperatures. Therefore, T_i^s and T_i^o should be compared before deciding which of these variables is used.

Figure 4.1 shows the scatter plots of summer daily T_i^s against T_i^o for the grid box covering the meteorological stations in London and Budapest (refer to Fig. 3.5) from 1991 to 2005. For both locations, T^s and T^o are strongly correlated with Pearson product-moment correlation coefficients of $\rho = 0.949$ and $\rho = 0.985$ for London and Budapest respectively. Consider the linear models

$$T_i^s = \beta_0 + \beta_1 T_i^o + \epsilon_i, \quad (4.4)$$

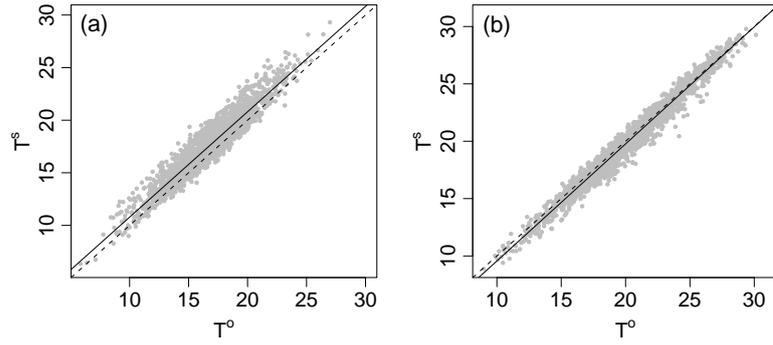


Figure 4.1: Scatter plots of summer T_i^s against T_i^o (in $^{\circ}\text{C}$) for the grid box covering the meteorological stations in (a) London and (b) Budapest from 1991 to 2005. The solid line shows the linear model fit (4.4) for the data, while the dashed line is the line of equal values.

Location	$\hat{\beta}_0$ ($^{\circ}\text{C}$)	$\hat{\beta}_1$	R^2
London	0.93 (0.12)	0.997 (0.008)	0.90
Budapest	-0.70 (0.08)	1.024 (0.004)	0.97

Table 4.1: Parameter estimates and R-squared statistics (proportion of variation explained) of model (4.4) for London and Budapest. Standard errors are given in parentheses. All the parameter estimates are significantly different from zero at the 5% level.

where β_0 and β_1 are the model parameters and ϵ represents Normal error with zero mean and constant variance. For London, a fitted positive intercept of $\hat{\beta}_0 = 0.93$ $^{\circ}\text{C}$ (standard error 0.12 $^{\circ}\text{C}$) together with a slope $\hat{\beta}_1$ close to unity (Table 4.1) indicate that station mean temperatures T_i^s are on average about 0.9 $^{\circ}\text{C}$ higher than the gridded mean temperatures T_i^o , as shown by the solid line in Fig. 4.1(a). This can possibly be explained by the location where T^s is measured. As the London Weather Centre station is situated in a built-up area in central London, temperatures observed at the station are likely higher than its surrounding areas, and are therefore systematically higher than T_i^o . Jones and Lister (2009) estimated the magnitude of urban heat island (the difference between urban and rural temperatures) for monthly mean temperatures at London Weather Centre to be 1.8 $^{\circ}\text{C}$ in the period 1981 to 2006. For Budapest, where the station is located in the suburban areas,

there are little systematic differences between T_i^s and T_i^o as the fitted line in Fig. 4.1(b) is close to the diagonal.

The difference between extreme T_i^s and T_i^o can be examined using the return level plots in Fig. 4.2. A return level refers to the quantile expected to be exceeded with a probability which is expressed in terms of the return period. The return curves are estimated by fitting generalized Pareto (GP) distributions to excesses of T_i^s and T_i^o above thresholds, which are chosen to be the respective sample 0.95 quantiles. Details on the GP distribution and return level plots are explained in Section 5.3.3 of Chapter 5 where such plots are used more extensively, and also in Appendix C. Comparing the return curves for T_i^s and T_i^o in London [Fig. 4.2(a)], the difference between extreme values of T_i^s and T_i^o appears to be larger (about 2 °C) than the difference predicted by the linear model. However, considering the sampling uncertainties indicated by the 90% confidence intervals (dashed lines), there is no strong evidence of significant underestimation of extreme T_i^o . For Budapest [Fig. 4.2(b)], there are little differences between high values of T_i^s and T_i^o . In Section 4.5, mortality models with T_i^o and T_i^s will be compared to confirm whether T_i^o is a more suitable temperature covariate.

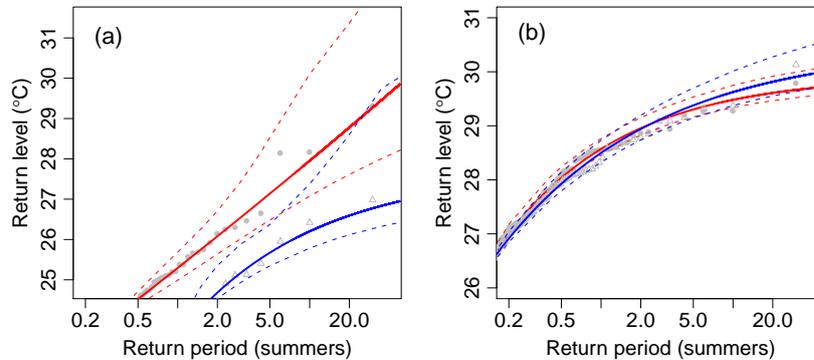


Figure 4.2: Return level plots of summer T_i^s and T_i^o for (a) London and (b) Budapest. The return curves for T_i^s and T_i^o , estimated by GP distribution fits, are shown by red and blue solid lines respectively. The corresponding 90% confidence intervals are shown by dashed lines. The empirical estimates of return levels for T_i^s and T_i^o are shown by grey dots and triangles respectively.

4.2.4 Modelling the functional dependence on covariates

At extratropical locations, the mortality annual cycles appear to be negatively associated with the air temperature annual cycles (see for example Figs 3.1 and 3.2). However, most modelling studies do not explain the mortality annual cycles solely by changes in air temperature because part of the mortality seasonality is believed to be related to other factors, such as seasonal changes in lifestyle, diet, stress and blood pressure (Donaldson et al., 2002; Huynen et al., 2001). A number of methods have been used to control for this mortality seasonality. Curriero et al. (2002), Gouveia et al. (2003), Kan et al. (2007), Kovats et al. (2004) and O'Neill et al. (2003) included a smoothing function of time estimated by smoothing splines with different degrees of freedom (normally 6 to 12 per year) in their mortality models. This method is commonly used in observational studies where only the relationship between mortality and environmental variables is of interest. The mortality seasonality which is modelled by this method does not have a fixed annual cycle for different years. This is an appropriate strategy for this study because as noted in Section 3.2.2, the amplitudes of mortality annual cycles in Budapest appear to decrease with time. However, when a smoothing function is used, the choice of degrees of freedom (or flexibility) affects how much mortality variation is left to be explained by changes in air temperature and other potential environmental covariates (or other variables). In a simulation study, Peng et al. (2006) demonstrated how the biases in the estimates of the dependence of mortality on environmental covariates vary with the degree of smoothing on the mortality time series.

Smoothing on the mortality time series can be performed as a first step in order to estimate the 'baseline mortality' on each day. The 'excess mortality', which is the difference between the observed mortality count and the estimate of the baseline mortality on that day, then becomes the response variable of the models. Examples of this strategy include calculating running means of various lengths (Dessai, 2002; Gosling et al., 2007; Sheridan and Kalkstein, 2004) and using average mortality in the same period in preceding years (Johnson et al., 2005). The baseline mortality estimated by these methods may be easily affected by fluctuations in mortality, such as a peak caused by an epidemic which persists

at a timescale comparable to the length of the moving average window chosen. Other methods for calculating ‘excess mortality’ are reviewed in Gosling et al. (2009a).

There is considerable variation in approaches to model relationships between the intra-seasonal response and environmental covariates. As described in Section 2.4.3, the dependency of daily mortality on air temperature in subtropical and temperate regions can usually be described by a U-shaped or a J-shaped curve (Basu and Samet, 2002; Curriero et al., 2002), where mortality increases towards the two extreme ends of temperature. Some modelling studies specify a parametric function to model this dependence. For example, Kunst et al. (1993) identified the air temperature at which minimum daily mortality was observed. The data set was then divided into two groups accordingly. Two separate linear functions were fitted to model the mortality-temperature relationships, one for the higher temperature range and one for the lower temperature range. Saez et al. (1995) also fitted separate linear functions for the two temperature ranges, but divided the data set into ‘summer’ and ‘winter’ according to the day of year. Laaidi et al. (2006) fitted a fourth-order polynomial for the whole temperature range. A possible problem of fitting a high order polynomial is the collinearity between the covariates which might lead to unreliable estimates of model parameters (Chapter 16 in Draper and Smith, 1998). In addition, as such parametric models are non-local fits, the mortality observations at the intermediate range of temperatures can affect the model fit at extreme temperatures which is more of interest. On the other hand, Kan et al. (2007) and Pauli and Rizzi (2008) modelled relationships between the response and each environmental covariate as non-linear smooth functions using smoothing splines. This again requires subjective specification of degrees of freedom of the function. Some observational studies (e.g. Curriero et al., 2002; Patten-den et al., 2003) first used a fully non-parametric GAM on their data sets to present the results graphically. A semi-parametric model, where the dependency of mortality on air temperature is modelled in a parametric form, while the dependency on other covariates is modelled by non-parametric smooth functions, was then fitted to the data again. This strategy provides quantitative estimates of changes in predicted mortality per unit change in air temperature.

4.3 Model evaluation methodologies

Model evaluation is an essential part of statistical modelling. The modeller should not only check whether the model is *adequate*, i.e. fits existing and out-of-sample data well, but also whether it is *well-specified*, i.e. is based on valid assumptions. Among the mortality modelling studies we have reviewed so far, some presented the evaluation work on model adequacy. For example, Dessai (2002), Gosling et al. (2007) and Laaidi et al. (2006) reported the R-squared or adjusted R-squared statistics for each model they fitted. These statistics provide basic indications of the goodness-of-fit of their models. Gosling et al. (2007) also performed split-sample validation where the model is only fitted to half of the data set and is then used to predict the response for the remaining half of the sample. The predictions are compared with the actual daily observations in terms of the correlation coefficient. For evaluation of model specification (validity of model assumptions), it is not routine practice to present such work, especially in observational epidemiological studies where the description of modelling results only involves model parameters and their uncertainties. It is then difficult for the reader to judge whether the models proposed in such studies are both adequate and well-specified. To remedy this problem, this section proposes a list of procedures to evaluate the adequacy and specification of mortality models. These procedures will be used systematically when comparing the models in the following section.

4.3.1 Model adequacy

The choice of model adequacy measures should be made according to how the statistical model is to be used. In this study, the prediction of extreme heat-related mortalities is the main area of concern. Apart from the accuracy in predicting daily changes in mortality, the ability of the models in predicting the upper tail of the mortality distribution is also important. Here are the measures that are used to assess model adequacy covering both aspects:

- (a) Un-biased risk estimator (UBRE; Wahba, 1990)

The goodness of fit of different mortality GAMs can be compared using UBRE:

$$\text{UBRE} = \frac{1}{n_o} D(\hat{\beta}) + \frac{2n_p}{n_o}$$

where $D(\hat{\beta})$ is the deviance of the fitted model, a measure related to the log-likelihood of the model (see Appendix B for details), n_p is the number of degrees of freedom in the model and n_o is the number of observations. This is a measure which focuses on daily predictions of mortality. A model with a lower UBRE score is preferred.

(b) Overdispersion statistic

The following statistic can be used to compare the adequacy of the systematic components of different models in explaining the mortality dispersion:

$$\hat{\phi} = \frac{D(\hat{\beta})}{n_r}$$

where n_r is the model's residual degrees of freedom. This statistic is an estimator for the scale parameter ϕ of the exponential family distribution (see Appendix B). For a Poisson GAM, a better model should have this statistic closer to 1.

(c) Mean residual life plot (Chapter 4 in Coles, 2001) for observed and predicted extreme mortality counts

The ability for a model to predict extreme heat-related mortalities can be assessed by comparing the upper tails of the distributions of observed and predicted daily mortality counts. The use of peak-over-threshold models from the classical extreme value theory to model such tails is not applicable for a Poisson distribution, the assumed distribution for mortality counts (Chapter 3 in Embrechts et al., 1997). Here the behaviour of the upper tail of mortality distribution is described by considering the quantity *mean excess* above a threshold u , $\mathbb{E}(Y - u | Y > u)$. This quantity represents the mean of the mortality excesses above u and can be estimated empirically using the sample mean. The change in the mean excesses of observed mortality y with different thresholds can be shown by a plot of locus of the points

$$\left(u, \frac{1}{n_u} \sum_{k=1}^{n_u} (y_{(k)} - u) \right)$$

where n_u is the number of observations above u and $\{y_{(1)}, y_{(2)}, \dots, y_{(n_u)}\}$ represent the observations which exceed u . To examine the differences between the upper tails of the

observed and predicted daily mortality distributions, this plot can be compared with a similar plot for the mean excesses of model predicted mortality counts $\mathbb{E}(\hat{\mu} - u | \hat{\mu} > u)$ which can also be estimated using the sample mean.

4.3.2 Model specification

The use of a Poisson GAM assumes the response conditional on the covariates $Y_i | \mathbf{X}_i$ to be independent and follow a Poisson distribution (4.1) with mean equal to the variance. The validity of these assumptions are assessed by considering the overdispersion statistic described in Section 4.3.1 and by examining the residual diagnostics and the coverage of prediction intervals.

(a) Residual diagnostics

If a model is well-specified, the deviance residuals ϵ^d (refer to Appendix B for details) are expected to be normally distributed with zero mean and variance of unity under sufficiently large samples, i.e. $\epsilon^d \sim N(0, 1)$. A number of residual diagnostic plots can therefore be used to assess the model specification. Firstly, a plot of deviance residuals against the model predicted values on the linear predictor scale should show an even scatter above and below zero with no obvious structure. This ensures that the response is independent. Secondly, the square root of the standardized deviance residuals against the model predicted values should be plotted (scale-location plot). If some structure is apparent from the plot, the constant variance approximation will be invalid. Thirdly, a normal quantile-quantile plot where the standardized deviance residuals are sorted and plotted against the standard normal distribution quantiles should be checked. For a well-specified model, a near straight line relationship should be observed in this plot.

(b) Confidence intervals and prediction intervals

In a regression model, the uncertainty in the predicted mean value of the response $\hat{\mu}$, given a set of explanatory variables, is associated with the uncertainty in the estimated model parameters. Confidence intervals of the predicted mean response can be constructed to represent such uncertainty. For example, in a mortality model, the meaning of 95% confidence intervals of predicted mean mortality, $[c_1, c_2]$, where $Pr(c_1 \leq \mu \leq c_2) = 0.95$,

is that we can be 95% confident that these intervals will contain the true value of mortality μ . For GAMs in this thesis which are fitted by the ‘mgcv’ package in R (Chapter 5 in Wood, 2006; R Development Core Team, 2009), the confidence intervals are estimated by bootstrapping methods (see Chapter 4.8 in Wood, 2006).

For practical uses of mortality models, the uncertainty in an individual predicted value of the response is more of interest. In addition to the model parameter uncertainty, an additional uncertainty arises from the variability of the assumed distribution of the response in the random component of the model. These uncertainties can be represented by constructing prediction intervals. For a Poisson GAM, 95% prediction intervals of each individual prediction, $[q_1, q_2]$, where $Pr(q_1 \leq y_+ \leq q_2) = 0.95$, are given by

$$\begin{aligned} q_1 &= Q_{0.025, \hat{\mu}} - (\hat{\mu} - c_1) \\ q_2 &= Q_{0.975, \hat{\mu}} - (c_2 - \hat{\mu}). \end{aligned}$$

In the above, $Q_{\lambda, \mu}$ is the $(100 \times \lambda)^{\text{th}}$ percentile of a Poisson distribution with the mean rate parameter μ . If the distributional assumption on the response is justified, given a set of explanatory variables, we can be 95% confident that a new individual observation y_+ will lie within the interval. The coverage of the prediction intervals on the existing data sample can be another diagnostic to assess the mortality model specification. For a well-specified model with a large enough data sample size, there should be about 2.5% of observations lying below the lower bound of the 95% prediction intervals ($y < q_1$) and about 2.5% lying above the upper bound of these intervals ($y > q_2$).

4.4 Model description

Based on the modelling strategies discussed in Section 4.2, a number of candidate statistical mortality models for London and Budapest are tested. By comparing the seven models described below, the following questions will be answered.

1. Does including humidity or using a biometeorological index improve mortality model predictions?

2. Is a model which specifies a non-parametric mortality-temperature relationship better than the model where such a relationship is specified to be parametric?
3. How do models using station observed temperatures T^o differ from models using gridded observed temperatures T^s ?

For the random components of the following mortality GAMs, the daily elderly mortality counts in each city Y_i given the explanatory variables \mathbf{X}_i are assumed to follow a Poisson distribution (4.1):

$$Y_i | \mathbf{X}_i \sim \text{Poisson}(\mu_i).$$

The systematic component of all the models include an offset $\log P_i$ to account for the changes in elderly population. To account for non-demographic and non-weather related variations in mortality, a smooth function with a predictor r_i is included. This variable represents time in years on a scale of $[0, 11]$ for London and $[0, 10]$ for Budapest, for example in the London series, for 1st January 1993 (the 1st day of the series), $r_1 = 0.0$; for 1st January 1997 (the 1462th day of the series), $r_{1462} = 4.0$. The smoothing functions are estimated by penalized cubic regression splines using the ‘mgcv’ package in R (Chapter 5 in Wood, 2006). Since only heat-related mortality is our concern in this study, the models are only fitted to data in the summer, defined as a four-month period from 15th May to 15th September each year. This choice is a compromise between having more observations in the analysis and avoiding days with low air temperatures which might affect the model fits.

- **Model BL:** This model estimates the baseline mortality by using only an offset for population plus a smooth function of r_i .

$$\log \mu_i = \log P_i + f_1(r_i)$$

The basis dimension (denoted by m ; see Appendix B for details) of f_1 , which can be considered as its upper limit of the degrees of freedom (d.f.), is chosen to be $m = 44$ for London and $m = 40$ for Budapest (same for the other models below). This represents a maximum of 1 d.f. for each month to ensure that weather-related variations of mortality are not captured.

- **Model TS:** In this model, the effect of daily mean temperature recorded at meteorological stations is represented by a smooth function of T_i^s .

$$\log \mu_i = \log P_i + f_1(r_i) + f_2(T_i^s) \quad (4.5)$$

To allow enough flexibility in the function representing the mortality-temperature relationship (f_2), the basis dimension is chosen such that the effective (actual) degrees of freedom (e.d.f.) is much lower than the ‘upper limit’ imposed (Section 4.1 in Wood, 2006; see also Appendix B). By trial and error, the basis dimension for f_2 is chosen to be $m = 10$.

- **Model TSH:** As in model TS, but a smooth function of daily mean vapour pressure e_i^s is added to investigate the humidity effect on mortality.

$$\log \mu_i = \log P_i + f_1(r_i) + f_2(T_i^s) + f_3(e_i^s)$$

By the same approach used to determine m for f_2 , the basis dimension for f_3 is chosen to be $m = 8$.

- **Model TAP:** Instead of air temperature, this model uses the daily mean value of ‘shade apparent temperature’ (Steadman, 1984), denoted by T_i^{ap} , as a covariate. For each location, sub-daily T^{ap} values are first calculated using $T^{ap} = -2.7 + 1.04T^s + 0.2e^s - 0.65v^s$, where v^s is the wind speed in m s^{-1} . The daily mean T_i^{ap} is then obtained by averaging the sub-daily values. This model investigates whether it is better to use a biometeorological index as the environmental covariate.

$$\log \mu_i = \log P_i + f_1(r_i) + f_4(T_i^{ap})$$

The basis dimension for f_4 is chosen to be $m = 10$.

- **Model PTS:** This model specifies the effects of air temperature on mortality in a parametric form with a second order polynomial of T_i^s .

$$\log \mu_i = \log P_i + f_1(r_i) + \beta_1 T_i^s + \beta_2 (T_i^s)^2$$

- **Model TO:** As in model TS, but gridded observed temperatures T_i^o are used as the covariate instead.

$$\log \mu_i = \log P_i + f_1(r_i) + f_5(T_i^o)$$

The basis dimension for f_5 is also chosen to be $m = 10$.

- **Model PTO:** As in model PTS, but T_i^o is used as the covariate.

$$\log \mu_i = \log P_i + f_1(r_i) + \beta_3 T_i^o + \beta_4 (T_i^o)^2$$

The above models consider only the effects of weather observed on the same day when mortality occurs. The appropriateness of this choice will be evaluated by examining the plots of autocorrelation function of the deviance residuals.

For the sake of brevity, some other tested models are not presented. These include the ones similar to model TS but use daily maximum or minimum air temperatures as covariates. Using the evaluation criteria for model adequacy described in Section 4.3.1, such models do not have better predictive performance compared to model TS. In addition, models which specifies the relationship between mortalities and temperatures in a parametric form using different orders of polynomials were tested. It was found that the model with a second order polynomial (model PTS) is better specified than the others.

4.5 Model evaluation and comparison

This section first describes the fit of model TS which uses station air temperature as the only environmental covariate, examines its residual diagnostics and compares it with the ‘baseline’ model BL. The questions listed in the last section are then addressed.

4.5.1 Model with air temperature effect

The time series of observed and predicted elderly mortality counts by model TS for London and Budapest are shown in Figs 4.3(a) and (c) respectively. Only the last three summers of the series are shown here for easier inspection. Compared to the baseline mortalities

predicted by model BL for the same period shown in Figs 4.3(b) and (d), model TS captures temperature-related spikes in mortality, most notably the one in London during the heatwave in August 2003 when the daily mean air temperature remained above 25 °C for two consecutive days. The 95% confidence intervals for the mean response and the 95% prediction intervals for individual observations, which are estimated as described in Section 4.3.2, are also shown in Fig. 4.3 to indicate the uncertainties of the predictions. The width of confidence intervals, which depend on the uncertainties of the model parameters, are much narrower than the prediction intervals, which also incorporate the uncertainties due to the inherent daily variability within the assumed response distribution. This shows that even if the model parameter uncertainty is low, there will be an intrinsic limit to the predictability of daily mortality simply because it is assumed to be a Poisson process.

The exponential transformed estimated smoothing function of station air temperature in model TS for both locations, $\exp(\hat{f}_2(T_i^s))$, are shown in Figs 4.4(a) and (b). Applying the exponential function on both sides of the model formula (4.5), it can be seen that $\exp(\hat{f}_2(T_i^s))$ represents the relative factor change in the proportion of elderly population dying, μ_i/P_i , as a result of the change in T_i^s for fixed r_i . This quantity is commonly known as ‘relative risk’ in epidemiology (Chapter 5 in Clayton and Hills, 1998). For example, Fig. 4.4(a) shows that elderly mortality risk in London estimated by model TS is about 70% higher at $T_i^s = 30$ °C compared to $T_i^s = 17$ °C. For both locations, the dependency of daily elderly mortality on daily mean air temperature is non-linear, with the mortality risk increasing more rapidly as T_i^s rises above 22 °C in London and above 25 °C in Budapest. The steeper curve for London suggests that elderly mortality in London is more sensitive to high air temperatures compared to Budapest.

The three plots on the top row of Fig. 4.5 are the residual diagnostics of model TS for London. The normal quantile-quantile plot [Fig. 4.5(a)] of deviance residuals shows roughly a straight line. No obvious structure can be observed in the plot of deviance residuals against the fitted value [Fig. 4.5(b)] and in the scale-location plot [Fig. 4.5(c)]. However, a number of outliers, which are large deviance residuals, can be observed in all three plots. These residuals, marked in red, relate to the data observed from 7th to 10th August during the 2003 heatwave. Since these observations are possibly influential

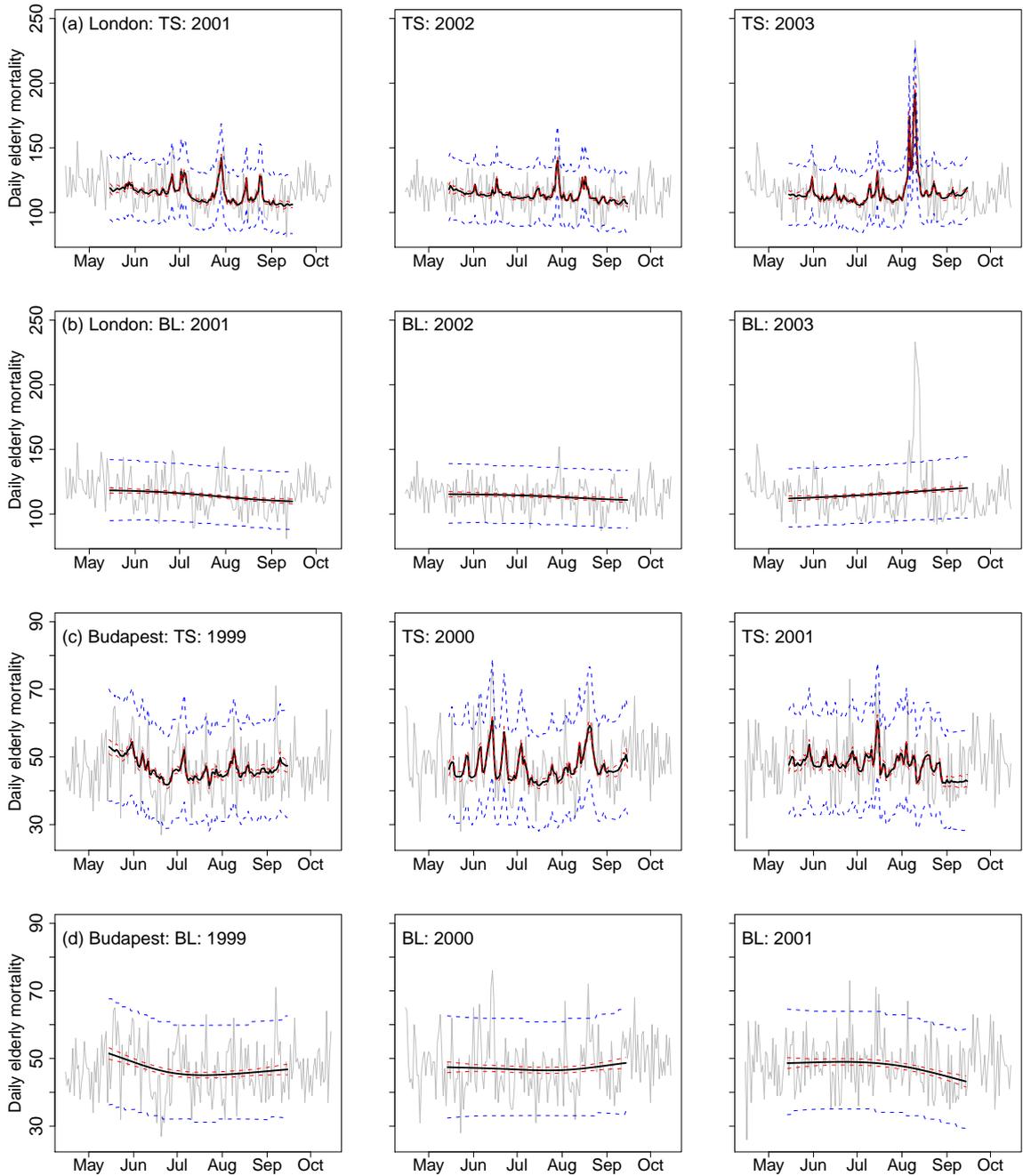


Figure 4.3: Time series of observed (grey) and predicted mortality counts (black) by models TS and BL for summers 2001 to 2003 in London (a, b) and summers 1999 to 2001 in Budapest (c, d). The 95% confidence intervals of the mean response and 95% prediction intervals of a new individual observation are indicated by red and blue dashed lines respectively.

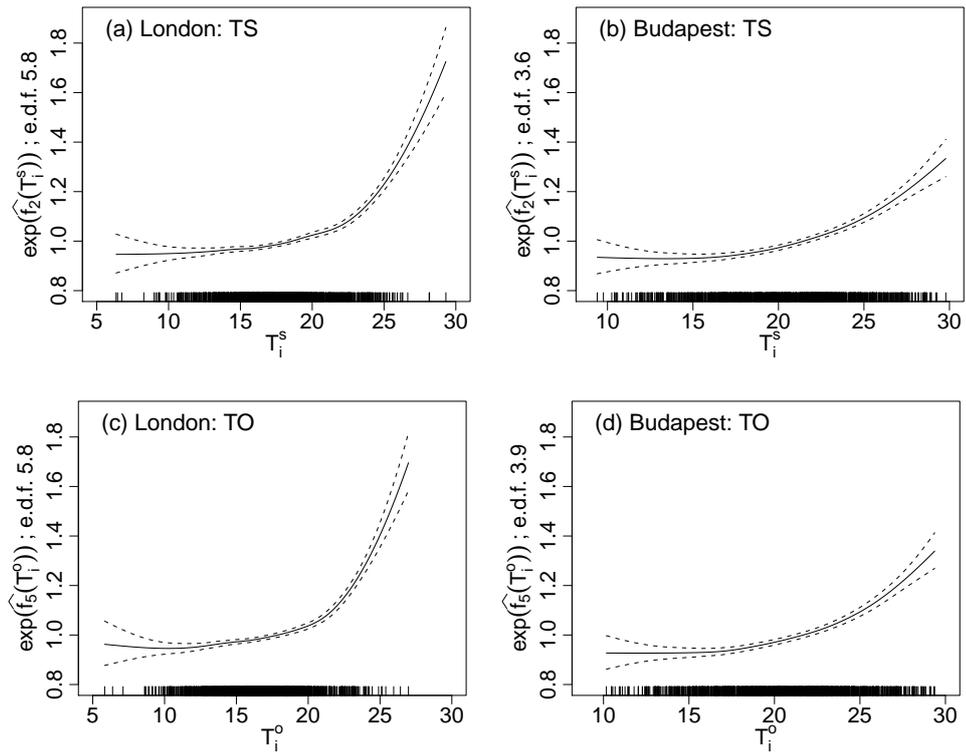


Figure 4.4: Exponential transformed estimated smoothing function of daily mean station T_i^s (in $^{\circ}\text{C}$) and gridded T_i^o (in $^{\circ}\text{C}$) air temperature in models TS ($\exp(\hat{f}_2(T_i^s))$; a and b) and TO ($\exp(\hat{f}_5(T_i^o))$; c and d) respectively for London and Budapest. The dashed lines are pointwise 2 standard error bands above and below the the estimate of each smoothing function. The densities of T_i^s and T_i^o observations are displayed by the rug plots at the bottom of each plot. The effective degrees of freedom (e.d.f.) of the estimated smoothing functions are shown on the y-axis.

to the model fit, the model will be fitted without them to observe any differences. The results will be discussed in Section 4.5.3. As for the residual diagnostics of model TS for Budapest (the bottom row in Fig. 4.5), no obvious violation of the model assumptions is apparent.

Comparing the evaluation statistics of models BL and TS (Table 4.2), the overdispersion statistic for model TS is smaller compared to model BL for both locations, as the inclusion of the air temperature effect in model TS increases the variance in the predicted mortality. This statistic is reasonably close to 1 for model TS, suggesting that there is no

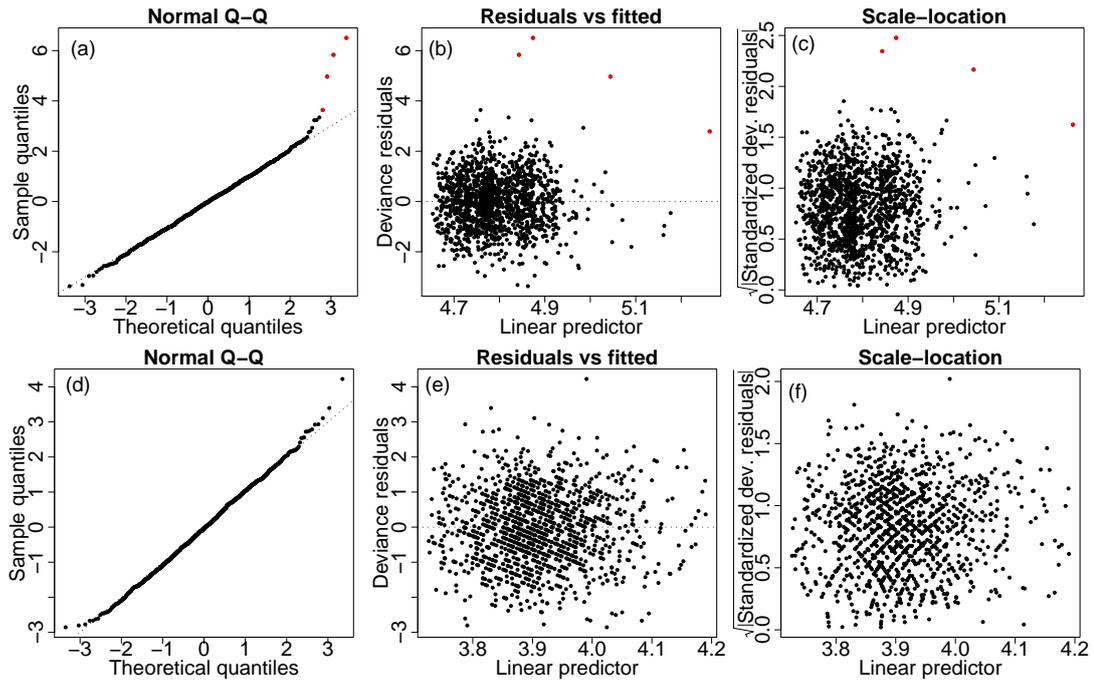


Figure 4.5: Residual diagnostic plots for model TS for London (top row) and Budapest (bottom row). The three panels on each row are: normal quantile-quantile plot of deviance residuals (a, d); deviance residuals against the fitted value on the linear predictor scale (b, e); square root of standardized deviance residuals against the fitted value on the linear predictor scale (c, f). The residuals for London from 7th to 10th August 2003 are highlighted in red.

large overdispersion relative to the assumed Poisson model. As for the 95% prediction interval coverage statistics, one should expect 2.5% of the observations lying below the lower bound ($y < q_1$) and 2.5% above ($y > q_2$) the upper bound of the intervals, i.e. about 34 observations on each side for London and 31 observations on each side for Budapest. For model BL, the numbers of $y > q_2$ are 40 and 39 for London and Budapest respectively. These indicate that the 95% prediction interval coverage for model BL is slightly too narrow because it does not take any effects of weather into account. However, the coverage for model TS appears slightly too wide, where only a total of about 4% and 3% of observations locate outside the interval for London and Budapest respectively.

(a) London				
Model	UBRE	Overdispersion	Prediction interval	
		$\hat{\phi}$	$y < q_1$	$y > q_2$
BL	0.507	1.667	38 (2.8%)	40 (2.9%)
TS	0.200	1.172	26 (1.9%)	29 (2.1%)
TSH	0.178	1.146	23 (1.8%)	24 (1.8%)
TAP	0.224	1.197	25 (1.9%)	29 (2.1%)
PTS	0.215	1.190	28 (2.1%)	29 (2.1%)
TO	0.183	1.162	24 (1.8%)	27 (2.0%)
PTO	0.210	1.191	24 (1.8%)	28 (2.1%)
(b) Budapest				
Model	UBRE	Overdispersion	Prediction interval	
		$\hat{\phi}$	$y < q_1$	$y > q_2$
BL	0.336	1.284	29 (2.3%)	39 (3.1%)
TS	0.149	1.121	19 (1.5%)	22 (1.8%)
TSH	0.143	1.114	18 (1.4%)	20 (1.6%)
TAP	0.162	1.134	21 (1.7%)	20 (1.6%)
PTS	0.147	1.120	19 (1.5%)	22 (1.8%)
TO	0.139	1.110	18 (1.4%)	20 (1.6%)
PTO	0.138	1.111	18 (1.4%)	22 (1.8%)

Table 4.2: Model evaluation statistics as described in Section 4.3 for the mortality models listed in Section 4.4. The meaning of symbols and acronyms are as follows. UBRE: un-biased risk estimator; $\hat{\phi}$: overdispersion statistic; $y < q_1$ and $y > q_2$: number and percentage of observations lying below and above the 95% prediction interval bounds of predicted elderly mortality counts.

4.5.2 The effect of humidity

The fits for the model with daily mean vapour pressure as a covariate (TSH) suggest that in addition to air temperature, there is a small additional effect of humidity in determining

elderly mortality at both locations. This is shown by the increasing relative risk with e_i^s in Fig. 4.6. The estimated smoothing functions of air temperature for this model are similar to that for model TS (not shown). For both locations, the mortality risk increases by not more than 10% as the vapour pressure reaches the highest present-day observed value, while the mortality risk increases more rapidly with air temperatures (see Fig. 4.4). The effects of air temperatures on mortality therefore dominate. For both cities, there are little differences between the time series of predicted elderly mortality counts by models TSH and TS (not shown), but the model evaluation statistics (Table 4.2) provide some evidence suggesting that TSH is a slightly better model. The UBRE score for model TSH is slightly lower than that of model TS at each location and is the lowest among the models which include T^s as a covariate.

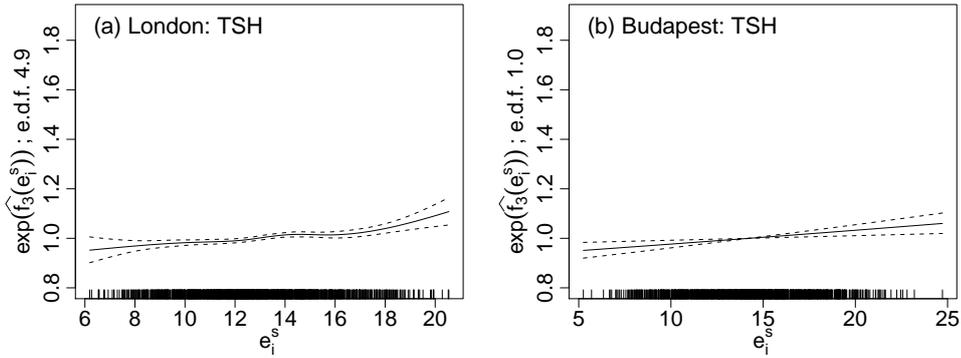


Figure 4.6: As in Fig. 4.4, but for the exponential transformed estimated smoothing functions of daily mean vapour pressure (e_i^s in hPa) in model TSH, $\exp(\hat{f}_3(e_i^s))$.

In terms of the model evaluation statistics, there is little evidence supporting the use of a model with daily mean apparent temperature as the covariate (TAP) over a model with daily mean air temperature (TS). For both locations, model TAP has higher UBRE scores (which mean worse fits) compared to both models TS and TSH. A comparison of time series plots for elderly mortality observed and predicted by models TAP and TS (not shown) reveals that the former model underestimates some of the mortality peaks.

As described in Section 4.2.2, a number of previous studies on the relationship between mortality and weather use daily apparent temperature as the covariate in their models, as this index has been considered to be a better indicator of the overall heat stress on humans

compared to air temperature. However, the results here suggest that using a model with this index could lead to worse predictive performance. In a study of mortality and meteorological data time series in three European cities, Hajat et al. (2006) also concluded that daily apparent temperature is a worse predictor of mortality compared to daily mean air temperature for London and Milan (Italy), but a better predictor for Budapest. The calculation of apparent temperature based on daily mean, rather than hourly air temperature, was cited as a possible reason for this result. There is no such problem in this study where the daily mean apparent temperature is computed based on hourly or 3-hourly measurements of air temperature, humidity and wind speed (see Section 4.4). In addition to model TAP, another model which includes an alternative biometeorological index, net effective temperature (Lee, 1980), was fitted to the London and Budapest data. The predictive performance of such a model was also found to be no better than model TS (results not shown). Biometeorological indices are therefore not chosen as covariates for the model used in Chapter 6 to project future heat-related mortalities in London and Budapest. Nevertheless, the above results do not mean that all biometeorological indices are necessarily worse predictors of mortality, nor these results necessarily apply to other cities.

For daily mean vapour pressure, the fits of model TSH indicate that this variable has a small effect in determining elderly mortality. Such a model slightly improves mortality predictions in terms of the UBRE scores. However, considering that the humidity effect is much smaller than the temperature effect, practically a model with only air temperature as the meteorological covariate should be adequate for projecting future mortalities. The daily mean vapour pressure is therefore not considered in the model for heat-related mortality projections.

4.5.3 Specification of the dependence of mortality on air temperature

For London, the fits of model TS (which specifies the dependence of mortality on temperature by a non-parametric smooth function) and PTS (which specifies such dependence by a second order polynomial) are noticeably different. The thick solid lines in Figs 4.7(a) and (b) show the predicted elderly mortality counts in London for the two models, as a

function of air temperature, while keeping the values of other covariates in the model fixed at their mean. Mortality predicted by model TS increases more rapidly compared to model PTS for T_i^s above 24 °C. As a test to investigate how these model fits are affected by the observations during the heatwave in 2003, both models are re-fitted with the year 2003 [black dots in Figs 4.7(a) and (b)] being omitted. The results are shown by the thick dashed lines. After removing the data for this year, models TS and PTS have similar predictions for the range of present-day observed air temperatures. Comparing the model predictions for each model indicated by thick solid and dashed lines on each panel, it appears that fully non-parametric model TS is less robust to the summer 2003 outliers. In contrast, comparing Figs 4.7(c) and (d), the mortality predictions for Budapest by the two models are similar.

The difference between the model fits of TS and PTS for London due to these outliers can explain their different overdispersion statistics (Table 4.2). As model TS predicts a higher number of elderly deaths at high air temperatures, the overdispersion of mortality captured by the model is greater than that captured by the model PTS, therefore the overdispersion statistic of model TS is closer to 1. The UBRE score of model TS is also lower, indicating that it has a better fit. For Budapest where the model fits by the models TS and PTS are similar, the overdispersion statistics and UBRE scores of the two models are similar.

We now focus on the predictions of the upper tails of the mortality distributions, using the mean residual life plots introduced in Section 4.3.1. The mean residual life plots of observed mortalities for London and Budapest are shown in Figs 4.8(a) and (d) respectively. The shapes of the upper tail of mortality distributions are related to the gradient of the solid lines, with a more negative slope indicating that the tail drops off more quickly. For London [Figs 4.8(a)], the mean excess mortality decreases with lower thresholds, but increases above $u = 140$, suggesting that the mortality distribution has a heavy upper tail. This feature is absent for the plot without the data from 2003, as shown by the blue line. The heavy upper tail of the mortality distribution is therefore related to the large number of deaths during the 2003 heatwave. Comparing this with the mean residual life plots of mortalities predicted by models TS and PTS [Figs 4.8(b) and (c)], it appears that the upper

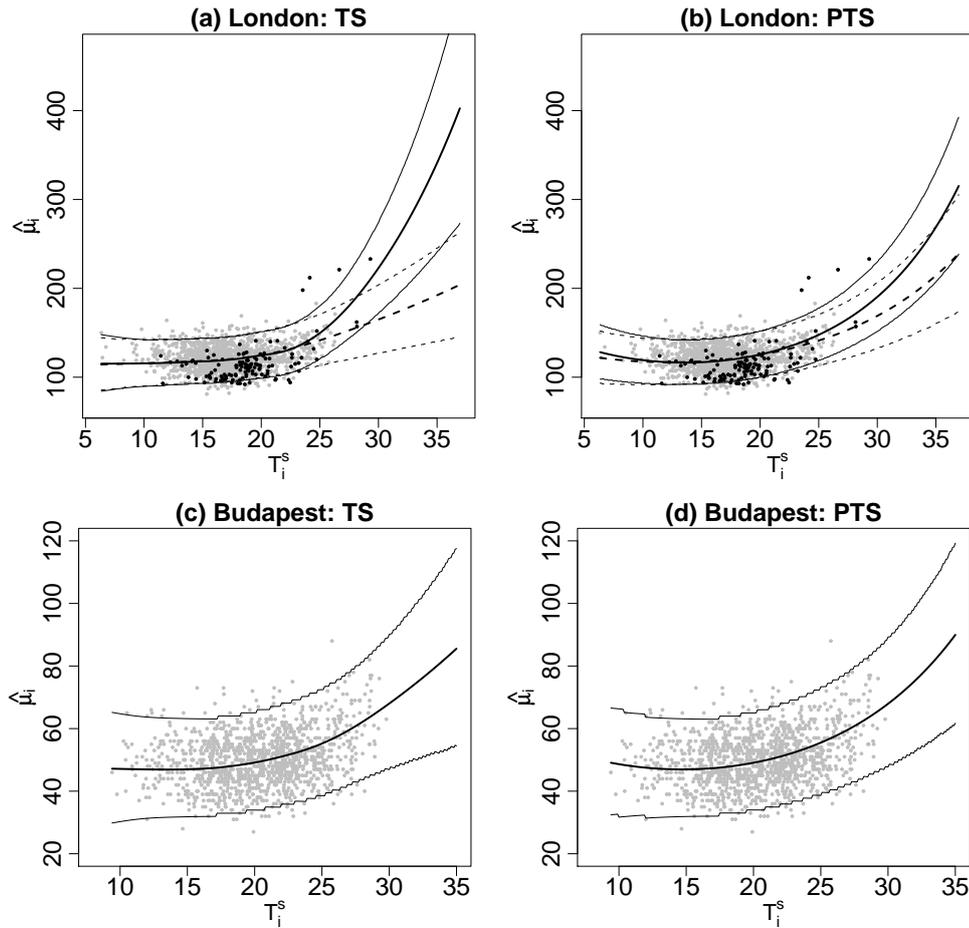


Figure 4.7: Predicted elderly mortality $\hat{\mu}_i$ (thick solid lines) by models TS and PTS as a function of daily mean station air temperature T_i^s (in °C) in London (a, b) and Budapest (c, d), with the values of population P_i and the time variable r_i in the models being kept fixed at their mean values. Observations are overlaid by grey dots, with those from 2003 in London being highlighted in black. The 95% prediction intervals are shown by thin solid lines. On panels (a) and (b), the thick dashed line represents the predictions for the same models fitted to all but 2003 summers. The thin dashed lines are the corresponding 95% prediction intervals.

tail of mortality is better predicted by model TS, albeit with large sampling uncertainties indicated by the 95% confidence intervals. For Budapest, the upper tails of the mortality distribution predicted by models TS and PTS [Figs 4.8(e) and (f)] have similar shapes and drop off more quickly compared to observed mortality distribution.

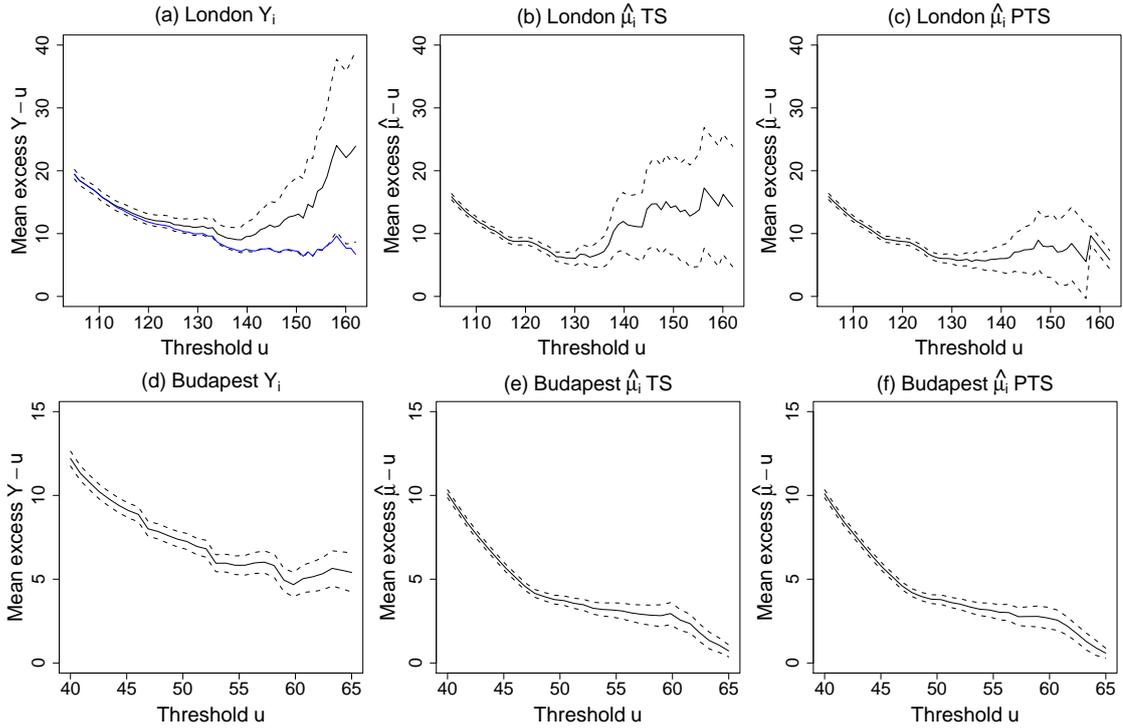


Figure 4.8: Mean residual life plots for observed (a, d) and predicted mortalities by models TS (b, e) and PTS (c, f) for London and Budapest. The 95% confidence intervals are indicated by dashed lines. The plot for observed mortality in London without the data in 2003 is superimposed on panel (a) by the blue line.

We now consider whether the fully non-parametric model or the model with parametric mortality-temperature relationship is more suitable for the projection of future mortalities later in this thesis. For Budapest, since the fits and the evaluation statistics of models TS and PTS are very similar, both models are equally suitable. For London, the choice is less obvious. Even though the evaluation statistics suggest that model TS fits the data better than model PTS, other factors need to be considered, for example how the different model sensitivities to outliers will affect the predictions. Under climate change, temperatures which are considered extremely high at the present day (say, those above the 99th percentile) are expected to become more common, and temperatures beyond the currently observed range are likely to occur in the future. When predicting future heat-related mortalities, an extrapolation of the estimated mortality-temperature relationship will be required. For model TS, where this relationship is represented in a non-parametric way, the

extrapolation depends on the shape of the estimated smoothing function at the extreme end of the range of observed temperature values. This is where the smoothing function can be heavily affected by outliers because of the lack of observations. As shown in Figs 4.7(a) and (b), at a mean air temperature of 35 °C (about 5 °C higher than the maximum value of mean air temperature observed in the sample), the number of daily elderly deaths in London predicted by model TS fitted with and without the data in 2003 differ by about 180, and their 95% prediction intervals do not overlap. In contrast, the corresponding difference for PTS is only about 70. Although in general, a model which is more robust to outliers is preferred, it is entirely possible that heat-related mortality increases more rapidly at extreme high air temperatures as predicted by model TS.

It is also important to note that even for model TS, where a non-parametric mortality-temperature relationship is specified, the high number of deaths occurred during the 2003 heatwave is still not well predicted, as shown by the large residuals in Figs 4.5(a) to (c). There are a number of possible reasons for the anomalously high mortalities recorded during this event. Firstly, the effect of a day with extremely high temperature on mortality may persist for a longer period of time. However, the plot of autocorrelation function of deviance residuals [Fig. 4.9(a)] shows that the autocorrelation is close to zero up to lags of 40 days. This lack of autocorrelation does not support the existence of persistent effect of extreme heat on mortality. Secondly, part of the elevated mortality during the event might be attributed to increased concentration of air pollution. Stedman (2004) estimates that 21 to 38% of excess mortality in the UK during the first two weeks of August 2003 are related to the increased concentration of ambient ozone and particulate matters. As explained in Section 4.2.2, such an effect is not considered in this study. Thirdly, sustained high temperatures during the heatwave might cause the number of deaths to be significantly higher than that predicted by mortality models as the heat stress on humans could not be relieved for a long period of time. Hajat et al. (2006) compares the observed summer mortality in London from 1976 to 2003 and the mortality predicted by a model which specifies the mortality-temperature relationship to be a linear function above temperatures of 20.5 °C. It was found that the observed mortalities during periods when daily mean temperatures are above the 99th percentile for 2 or more consecutive days are 5.5% above

that predicted by the model. Here in this study, since the time series of mortality data is relatively short, it is not possible to confirm whether there is such a ‘heatwave effect’ or to reliably model such an effect.

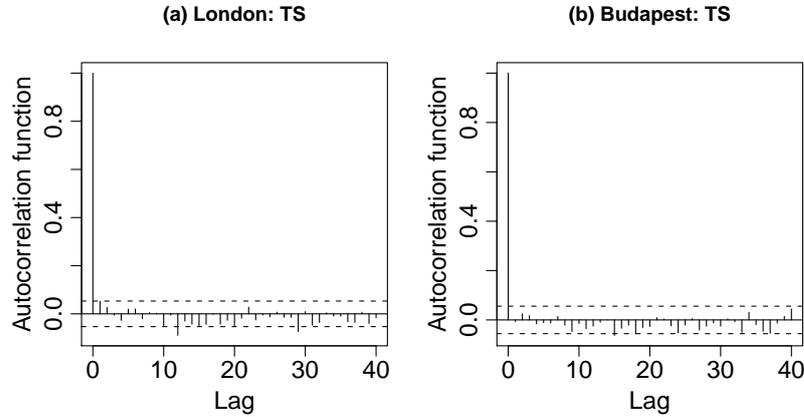


Figure 4.9: Autocorrelation function of deviance residuals of model TS for (a) London and (b) Budapest up to lag of 40 days. The dashed lines are 95% confidence intervals for zero autocorrelation.

Since the mortality predictions by model TS during the 2003 heatwave are more accurate than those of model PTS, the model used to project future heat-related mortalities will specify the mortality-temperature relationship with a non-parametric smoothing function. However, mortality projections using the model which specifies such a relationship with a second order polynomial will also be attempted as part of the sensitivity analysis in Chapter 6.

4.5.4 The use of gridded observed temperature as a covariate

There are minor differences in the estimated smoothing functions using station observed air temperatures (T_i^s) in mortality models and that using gridded observed air temperatures (T_i^o) in the models. The estimated smoothing functions of T_i^o in model TO for London and Budapest are shown in Figs 4.4(c) and (d) respectively. The shapes of the smoothing functions for each location are similar to the corresponding model TS [Figs 4.4(a) and (b)]. Compared to model TS, the estimated mortality-temperature curve of model TO

for London is shifted slightly to the left. This is consistent with the observation made in Section 4.2.3 that values of T_i^o are about 0.9 °C systematically lower than T_i^s .

The model evaluation statistics suggest that the use of gridded T_i^o gives a slightly better model fit. For both locations, model TO has lower UBRE scores than model TS. Model PTO, the ‘parametric’ counterpart of model TO, also has lower UBRE scores than model PTS for both locations. The improved goodness of fit of models using T_i^o provides some evidence that gridded observed temperatures better represent the general thermal conditions experienced by the population compared to the temperatures recorded at a single station. As a result, models with T_i^o as the temperature covariate, i.e. models TO and PTO, are chosen to project future mortalities.

In Sections 4.5.2 and 4.5.3, the discussion on the inclusion of humidity, prediction of the upper tails of mortality distributions and the sensitivity of models to outliers are based on mortality models using T_i^s . It should be noted that all the conclusions are also valid for the corresponding models using T_i^o because of the similarities between the fits of the models using the two variables.

4.6 Summary

In this chapter, the strategies of modelling the present-day heat-related mortality for London and Budapest have been discussed. Based on the comparison of the fits of a number of candidate models and the evaluation of their adequacy and specification, a Poisson GAM with gridded observed daily mean air temperature as the only meteorological covariate (model TO) is chosen as the main model for the projections of future heat-related mortalities in Chapter 6. In this model, the relationship between elderly mortality and temperatures is specified using a non-parametric smooth function. While this model is slightly less robust to outliers compared to the model which specifies the mortality-temperature relationship as a second order polynomial (model PTO), this model predicts the mean extreme excess mortality slightly better for London.

Chapter 5

Regional climate model temperature calibration

5.1 Aim

As discussed in Chapter 2, simulations of climate models contain discrepancies. When projecting future mortalities, the mortality models developed in the last chapter are not to be directly driven by the raw uncalibrated daily mean air temperature projections from HadRM3. This chapter describes the work on the calibration of HadRM3 summer temperature projections from 2010 to 2099. It starts with a discussion of two generic approaches for calibrating climate model variables, ‘bias correction’ and ‘change factor’. The effects of using these two approaches on the temperature projections of the HadRM3 perturbed physics ensemble, especially the extremes, are then compared for London and Budapest. At the end of this chapter, the calibrated temperatures of the standard version of HadRM3 for Europe are presented, in order to highlight the different results obtained by the two calibration approaches and to give an indication of how mortality risk will change in other European locations.

5.2 General approaches for calibration

Although this thesis focuses on the extreme HadRM3 summer temperature projections, this section first considers the more general case of calibrating the entire distribution of a climate variable simulated by climate models, not only limited to surface air temperatures. The application of the calibration methodologies discussed below on HadRM3 temperature projections will be described in Section 5.3.

5.2.1 Overview of approaches

The calibration of climate model projections involves four random variables (Fig. 5.1). For the present-day, we have both the actual (station or gridded) observations O and climate model simulations G . For the future, we have the model simulations of the climate variable, denoted by G' . The aim of the calibration exercise is to estimate the unknown future observed variable of interest, denoted by O' . This can be achieved by two distribution mapping approaches, ‘bias correction’ and ‘change factor’. For the purpose of calibration, the sequences of O and G in a specified ‘present-day’ time period and the sequences of O' and G' in a specified ‘future’ time period are each assumed to be independent and identically distributed. This assumption is appropriate for sufficiently short time periods where the trends in individual variables are sufficiently small. The cumulative distribution functions (c.d.f.) of these four variables are denoted by F_O , F_G , $F_{O'}$ and $F_{G'}$. In addition, when O' obtained using the bias correction approach and the change factor approach need to be distinguished, the symbols O'_b and O'_c will be used respectively.

The bias correction approach is based on the difference between the distributions of O and G , i.e. the bias in the present-day climate model simulations. This approach assumes that there exists a constant transfer function B such that the random variable G can be transformed to have the same distribution as the random variable O :

$$\begin{aligned}
 B(G) &\stackrel{d}{=} O \\
 \Leftrightarrow B \circ F_G^{-1} &= F_O^{-1} \\
 \Leftrightarrow B &= F_O^{-1} \circ F_G
 \end{aligned}$$

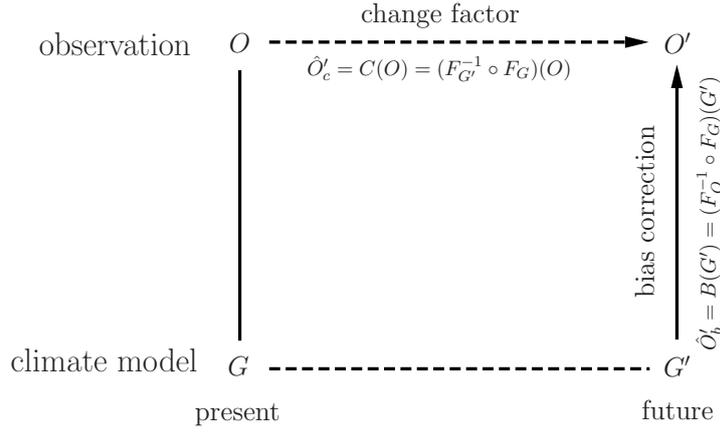


Figure 5.1: Schematic diagrams showing two approaches to calibrate projections of climate variables from climate models, bias correction (solid lines) and change factor (dashed lines).

where $\stackrel{d}{=}$ indicates equality in distribution, and F_O^{-1} and F_G^{-1} denote the inverse of F_O and the inverse of F_G respectively. If B is assumed to stay constant in the future, which means that model biases do not change with time, future model projections G' can then also be transformed to have the same distribution as O' :

$$\hat{O}'_b = B(G') = (F_O^{-1} \circ F_G)(G').$$

On the other hand, the change factor approach is based on the difference between the distributions of G and G' , i.e. the change in the distribution of the modelled variable with time. This approach assumes that there exists a different transfer function C such that G can be transformed to have the same distribution as G' :

$$\begin{aligned} C(G) &\stackrel{d}{=} G' \\ \Leftrightarrow C \circ F_G^{-1} &= F_{G'}^{-1} \\ \Leftrightarrow C &= F_{G'}^{-1} \circ F_G. \end{aligned}$$

By definition, C is an identity function if there are no future changes in the distribution of the climate model variable, i.e. $G' \stackrel{d}{=} G$. Assuming that the present-day observations O can be transformed by C such that they have the same distribution as O' ,

$$\hat{O}'_c = C(O) = (F_{G'}^{-1} \circ F_G)(O).$$

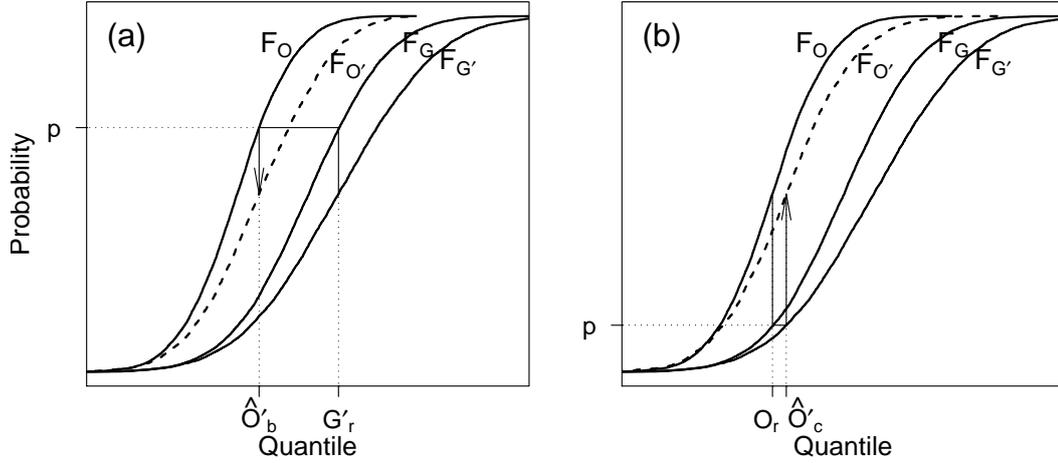


Figure 5.2: Illustration of distribution mapping by (a) bias correction and (b) change factor. The thick solid curves show three cumulative distribution functions (c.d.f.) F_O , F_G and $F_{G'}$ as labelled, while the thick dashed curve is the unknown c.d.f. $F_{O'}$ estimated by the two calibration methods. Examples of how values from the distributions of G' and O are mapped are shown by arrows.

The mapping by the transfer functions B and C is illustrated in Fig. 5.2. For bias correction [Fig. 5.2(a)], given a value G'_r from the distribution of G' , the probability $p = Pr(G \leq G'_r)$ is first obtained by $F_G(G'_r)$. The bias-corrected value O'_b is then estimated by $F_{O'}^{-1}(p)$ such that $p = Pr(O \leq O'_b)$. For change factor [Fig. 5.2(b)], given a value O_r from the distribution of O , $F_O(O_r)$ gives the probability $p = Pr(G \leq O_r)$. The calibrated value O'_c is then estimated by $\hat{O}'_c = F_{G'}^{-1}(p)$.

In practice where F_O , F_G and $F_{G'}$ need to be estimated, the form of B and C can be specified in a number of ways. A straightforward option is to estimate F_O , F_G and $F_{G'}$ by the corresponding empirical distribution functions (e.g. Ines and Hansen, 2006). However, there are two problems with this empirical quantile mapping method. The small number of observations at the tails of the O , G , G' distributions may cause the calibration to be less precise. In addition, it is not possible to perform out-of-sample calibration due to the change of support in O , G and G' . Both approaches can only map values in the domain of F_G , therefore for example, values of G' greater than the maximum value within the sample of G cannot be bias-corrected. Another possible choice is to fit theoretical distributions to

O , G , G' and then to perform the calibration based on the theoretical quantiles. For example, in the calibration of monthly GCM temperature output for the UK, Vidal and Wade (2008) fitted normal distributions to observed temperatures and GCM simulated present-day temperatures. The biases of GCM temperatures for the future are then corrected by mapping the quantiles of the fitted distributions. However, this method is not desirable if the actual distributions of observed or model variables deviate significantly from the assumed theoretical distributions.

This thesis considers a different approach to specify the transfer functions, by assuming certain relationships between the properties (location, scale and shape) of the distributions involved. Ferro et al. (2005) has used similar ideas to interpret model simulations of climate change.

5.2.2 Transfer functions for bias correction

For the bias correction approach, the transfer function B is specified by assuming relationships between O and G , i.e. the distributions of the observed and modelled climate variable in the present-day. The following three methods are considered.

(a) Bias correction in location (BC-L)

Assume that the distributions of O and G have different location, but have the same scale and shape, i.e. the RCM simulations have biases in the location only:

$$F_O(O) = F(O - \alpha_O) \quad (5.1)$$

$$F_G(G) = F(G - \alpha_G) \quad (5.2)$$

where α_O and α_G are the location parameters of O and G respectively. The above expressions mean that $(O - \alpha_O)$ and $(G - \alpha_G)$ belong to the same distribution whose c.d.f. is F .

The transfer function is then derived as follows:

$$\begin{aligned} \hat{O}'_b = B(G') &= (F_O^{-1} \circ F_G)(G') \\ &= F_O^{-1} [F(G' - \alpha_G)] \\ &= \alpha_O + F^{-1} [F(G' - \alpha_G)] \\ &= \alpha_O + (G' - \alpha_G), \end{aligned} \quad (5.3)$$

where the equality between the first and the second lines uses (5.2), and the equality between the second and the third lines uses (5.1).

(b) Bias correction in location and scale (BC-LS)

Assume that the distributions of O and G to have different location and scale, but the same shape:

$$\begin{aligned} F_O(O) &= F\left(\frac{O - \alpha_O}{\beta_O}\right) \\ F_G(G) &= F\left(\frac{G - \alpha_G}{\beta_G}\right) \end{aligned} \quad (5.4)$$

where β_O and β_G are the scale parameters of O and G respectively. Here the biases in the location are assumed to be additive and the biases in the scale are multiplicative. The corresponding transfer function can be derived in a similar manner as in case (a):

$$\begin{aligned} \hat{O}'_b = B(G') &= (F_O^{-1} \circ F_G)(G') \\ &= F_O^{-1}\left[F\left(\frac{G' - \alpha_G}{\beta_G}\right)\right] \\ &= \alpha_O + \beta_O F_O^{-1}\left[F\left(\frac{G' - \alpha_G}{\beta_G}\right)\right] \\ &= \alpha_O + \frac{\beta_O}{\beta_G}(G' - \alpha_G). \end{aligned} \quad (5.5)$$

This location and scale correction method was adopted by Leith (2007) to calibrate GCM projections of temperature, humidity and sea-level pressure, which are then used to statistically downscale (see Section 2.3.2) rainfall projections for multiple sites in the UK.

(c) Bias correction in location and scale for Box-Cox transformed variables (BC-LSB)

Assume that the location, scale and shape of the distributions of O and G are different. Their different shapes are accounted for using the Box-Cox transformations A_O and A_G which are applied on O and G respectively:

$$\tilde{O} = A_O(O) = \begin{cases} \frac{O^{\lambda_O} - 1}{\lambda_O} & \text{for } \lambda_O \neq 0 \\ \log(O) & \text{for } \lambda_O = 0 \end{cases} \quad (5.6)$$

$$\tilde{G} = A_G(G) = \begin{cases} \frac{G^{\lambda_G} - 1}{\lambda_G} & \text{for } \lambda_G \neq 0 \\ \log(G) & \text{for } \lambda_G = 0, \end{cases} \quad (5.7)$$

where $\tilde{(\cdot)}$ is used to indicate Box-Cox transformed variables. The transformation is only applicable when all values in the distribution are positive. The Box-Cox transformation attempts to modify a distribution such that it becomes close to a normal distribution, with suitable choices of transformation parameters λ_O and λ_G . For $\lambda < 1$, larger values in a sample are decreased more compared to the smaller values, making a positively skewed distribution closer to being symmetrical. For the reverse case, $\lambda > 1$ causes larger values in the sample to increase more compared to smaller values, therefore a negatively skewed distribution is made more symmetric (Chapter 3 in Wilks, 2006). Eastoe and Tawn (2009) applied the Box-Cox transformation to preprocess the non-stationarity in surface ozone data before modelling their extremes.

Assuming $\tilde{O} \sim N(\alpha_{\tilde{O}}, \beta_{\tilde{O}}^2)$ and $\tilde{G} \sim N(\alpha_{\tilde{G}}, \beta_{\tilde{G}}^2)$, i.e. the transformed variables \tilde{O} and \tilde{G} follow normal distributions with mean $\alpha_{\tilde{O}}$ and $\alpha_{\tilde{G}}$ and variance $\beta_{\tilde{O}}^2$ and $\beta_{\tilde{G}}^2$ respectively, then \tilde{O} and \tilde{G} can be assumed to be different in location and scale only:

$$F_{\tilde{O}}(\tilde{O}) = \Phi\left(\frac{\tilde{O} - \alpha_{\tilde{O}}}{\beta_{\tilde{O}}}\right)$$

$$F_{\tilde{G}}(\tilde{G}) = \Phi\left(\frac{\tilde{G} - \alpha_{\tilde{G}}}{\beta_{\tilde{G}}}\right),$$

where $F_{\tilde{O}}$ and $F_{\tilde{G}}$ are the c.d.f. of \tilde{O} and \tilde{G} respectively, and Φ is the c.d.f. of the standard normal distribution $N(0, 1)$. A correction similar to that in case (b) can then be applied on \tilde{G} with the transfer function

$$\hat{\tilde{O}}' = B(\tilde{G}') = \alpha_{\tilde{O}} + \frac{\beta_{\tilde{O}}}{\beta_{\tilde{G}}}(\tilde{G}' - \alpha_{\tilde{G}}), \quad (5.8)$$

where \tilde{O}' can be back-transformed to O' by the inverse of Box-Cox transformation A_O (5.6).

As a result, if the shapes of both O and G are assumed not to change in the future, such that $\tilde{G}' = A_G(G')$ and $\tilde{O}' = A_O(O')$ also follow normal distributions, the biases of climate model projections for the future G' can be corrected by

$$\hat{O}'_b = (A_O^{-1} \circ B \circ A_G)(G') \quad (5.9)$$

with B as specified in (5.8). For example, in the case where $\lambda_O, \lambda_G \neq 0$,

$$\hat{O}'_b = \left\{ 1 + \lambda_O \left[\alpha_{\hat{O}} + \frac{\beta_{\hat{O}}}{\beta_{\hat{G}}} \left(\frac{G^{\lambda_G} - 1}{\lambda_G} - \alpha_{\hat{G}} \right) \right] \right\}^{\frac{1}{\lambda_O}}.$$

5.2.3 Transfer functions for change factor

For the following two methods under the change factor approach, the transfer function C is specified by assuming relationships between G and G' , i.e. the distributions of the modelled climate variable in the present-day and in the future. The derivation of the transfer functions is similar to that shown in methods (a) and (b) of the bias correction approach in Section 5.2.2 and is therefore omitted below.

(a) Change factor in location (CF-L)

Assume that the distributions of G and G' are only different in their locations,

$$F_G(G) = F(G - \alpha_G)$$

$$F_{G'}(G') = F(G' - \alpha_{G'})$$

where $\alpha_{G'}$ is the location parameter of G' . The transfer function is given by

$$\hat{O}'_c = C(O) = \alpha_{G'} + (O - \alpha_G). \quad (5.10)$$

This technique is widely adopted in climate change impact studies, including the mortality projections performed by Donaldson et al. (2002) and Dessai (2003), which are reviewed in Section 2.5. In these projections, the model projected changes in mean temperatures are added to the sequence of present-day observed temperatures to obtain the future projected temperature sequence.

(b) Change factor in location and scale (CF-LS)

Assume that between the distributions of G and G' , there is an additive change in the location and a multiplicative change in the scale:

$$\begin{aligned} F_G(G) &= F\left(\frac{G - \alpha_G}{\beta_G}\right) \\ F_{G'}(G') &= F\left(\frac{G' - \alpha_{G'}}{\beta_{G'}}\right) \end{aligned} \quad (5.11)$$

where $\beta_{G'}$ is the scale parameter of G' . The corresponding transfer function is

$$\hat{O}'_c = C(O) = \alpha_{G'} + \frac{\beta_{G'}}{\beta_G}(O - \alpha_G). \quad (5.12)$$

A similar change factor in location and scale method is applied by UKCP09 to calibrate climate variables from HadRM3, which are then used to drive weather generators (see Section 2.3.2) to obtain regional climate projections (Jones et al., 2009).

A third change factor method analogous to method (c) of the bias correction approach could be attempted to account for changes between the shapes of the G and G' distributions, however such a case is not presented here. This is because the use of such technique needs to assume that the distributions of O and G have the same shape. As will be shown in Section 5.6, this assumption is considered not valid for the calibration of HadRM3 temperatures.

5.2.4 Comparison of approaches

The main features of the two calibration approaches presented above are compared in Table 5.1, but this comparison does not give a clear indication as to which approach is more reliable. Importantly, the distribution of calibrated temperatures produced by bias correction and change factor, O'_b and O'_c , can be different even when the assumptions involved in the specification of corresponding transfer functions are satisfied. Consider, for example, the bias correction in location and scale (BC-LS) (5.5) and change factor in location and scale (CF-LS) (5.12). The population mean of O'_b and O'_c are not necessarily identical:

$$\begin{aligned} \mathbb{E}(O'_b) &= \alpha_O + \frac{\beta_O}{\beta_G}(\alpha_{G'} - \alpha_G) \\ \mathbb{E}(O'_c) &= \alpha_{G'} + \frac{\beta_{G'}}{\beta_G}(\alpha_O - \alpha_G) \end{aligned} \quad (5.13)$$

$$\mathbb{E}(O'_b) - \mathbb{E}(O'_c) = \alpha_O - \alpha_{G'} + \frac{1}{\beta_G} [\beta_O (\alpha_{G'} - \alpha_G) - \beta_{G'} (\alpha_O - \alpha_G)], \quad (5.14)$$

while the population variance is the same:

$$\text{var}(O'_b) = \text{var}(O'_c) = \frac{\beta_O^2 \beta_{G'}^2}{\beta_G^2}. \quad (5.15)$$

This means that different results will be obtained from the two approaches even if the shapes are the same between O and G (5.4) and between G and G' (5.11).

Bias correction	Change factor
$B(G') = (F_O^{-1} \circ F_G)(G')$	$C(O) = (F_{G'}^{-1} \circ F_G)(O)$
<ul style="list-style-type: none"> • based on relationships between O and G • no assumptions on relationship between G and G' • assumed to be independent on time • performance can be evaluated by bias correction of G 	<ul style="list-style-type: none"> • based on relationships between G and G' • depends on time implicitly by differences between G and G' • performance cannot be evaluated since the change factor is identity for no change in time

Table 5.1: Comparison of the two model calibration approaches.

Many previous climate change impact studies adopt one particular calibration method under either the bias correction approach or the change factor approach to calibrate future climate variables from model output. In Sections 5.4 to 5.6, calibration using all five methods under the two calibration approaches will be performed, such that the impacts of the choice of methods can be compared. The reliability of different methods will be assessed by examining the validity of the distributional assumptions involved. In addition, the performance of different methods of the bias correction approach can also be evaluated by applying them on G , the model simulated climate variable for the present-day period.

Finally, it should be noted that there have been other proposed climate model calibration methods. One example is given in the mortality projections by Gosling et al. (2009b) discussed in Section 2.5. In their calibration of GCM projected temperatures, the location and scale parameters of O , G and G' were first estimated by fitting logistic distributions to

the data¹. Time series of O' to drive the mortality model is produced by random sampling of a logistic distribution with $\alpha_{O'} = \alpha_O + (\alpha_{G'} - \alpha_G)$ and $\beta_{O'} = \beta_O + (\beta_{G'} - \beta_G)$. The rationale behind the approach is somewhat similar to the change factor approach. However, the change factor in scale proposed in this section is multiplicative instead of additive, and it can be observed from (5.13) that the population mean of O'_c also depends on the scale parameters β_G and $\beta_{G'}$.

5.3 Calibration of HadRM3 temperatures

As mentioned in the previous section, this thesis attempts to calibrate the entire distribution of HadRM3 projected summer temperatures for each of its grid boxes using different approaches. The extremes of temperatures calibrated by different methods are then estimated and compared. For London and Budapest, temperature projections for each of the 11 HadRM3 PPE members are calibrated individually, while for other European locations, only the standard version of HadRM3 is considered. This section describes how the ideas discussed in the previous section are applied to perform the temperature calibration.

The variables involved in this calibration include the present-day E-OBS gridded daily mean air temperatures (T^o), HadRM3 simulated temperatures for the same period (T^g), future HadRM3 projected temperatures (denoted by $T^{g'}$) and calibrated temperatures (denoted by $T^{o'}$). These correspond to O , G , G' and O' respectively in the discussion of the previous section. For notational convenience, the latter set of symbols will be used to represent the temperature variables for the rest of this chapter. For the purpose of temperature calibration and estimation of extremes, the ‘present-day’ is defined to be the 30-year period from 1970 to 1999. As recommended by the World Meteorological Organization, it is common to define the ‘present climate’ using sample statistics of weather observations over 30 years (World Meteorological Organization, 1989; Räisänen and Ruokolainen, 2008). For the future, three 30-year future periods (time-slices) are considered: 2010 to

¹A different set of symbols are used in Gosling et al. (2009b) to represent the temperature variables and location and scale parameters, but symbols consistent with that used in Section 5.2 are used here to allow easier comparison between different calibration approaches.

2039, 2040 to 2069 and 2070 to 2099. The observed and HadRM3 simulated temperatures in each period are each assumed to be independent and identically distributed.

5.3.1 Estimation of parameters

For the three methods under the bias correction approach, HadRM3 projected temperatures from 2010 to 2099 are calibrated using (5.3), (5.5) and (5.9) respectively. HadRM3 simulated temperatures for the present-day (1970 to 1999) are also calibrated by replacing G' with G in these formulae. For BC-L and BC-LS, the location and scale parameters, i.e. α_O , α_G , β_O and β_G , are estimated by the sample mean and standard deviation respectively. It should be noted that alternative choices are available to estimate these parameters, such as the sample median and sample interquartile range. However, the results to be presented in Sections 5.4 to 5.6 are found to be generally insensitive to this choice. As for BC-LSB, constants γ_O and γ_G are first added to sample values of O and G in (5.6) and (5.7) before the calibration. This is to ensure that all the values in both samples are positive, such that Box-Cox transformation is valid. Take O as an example, (5.6) becomes

$$\tilde{O} = \begin{cases} \frac{(O + \gamma_O)^{\lambda_O} - 1}{\lambda_O} & \text{for } \lambda_O \neq 0 \\ \log(O + \gamma_O) & \text{for } \lambda_O = 0. \end{cases}$$

Here γ_O is given by

$$\gamma_O = \begin{cases} 0 & \text{if all } O_i > 0 \\ -\min(O_i) & \text{otherwise,} \end{cases}$$

where $\min(O_i)$ is the minimum value in the sample of O . \tilde{G} is defined by adding the constant γ_G in a similar manner. The parameters λ_O , $\alpha_{\tilde{O}}$, $\beta_{\tilde{O}}$, λ_G , $\alpha_{\tilde{G}}$ and $\beta_{\tilde{G}}$ are then estimated by maximum likelihood (Box and Cox, 1964).

For the two methods under the change factor approach, the parameters α_G and β_G in (5.10) and (5.12) are estimated by the mean and standard deviation of the sample of G , while $\alpha_{G'}$ and $\beta_{G'}$ for each of the three 30-year time-slices are estimated by the mean and standard deviation of the corresponding time-slices of G' . The calibrated temperatures O' for the three time-slices are then calculated using (5.10) and (5.12).

5.3.2 Comparison of temperature distributions

In order to assess the validity of distributional assumptions of different calibration methods, the distributions of O , G and G' need to be compared. For individual grid boxes (e.g. London and Budapest), pairwise comparison of different distributions is done by inspecting the sample quantile-quantile plots. For gridded data sets with many variables, this procedure is not practical, and so spatial maps of selected sample quantiles are compared. In this thesis the sample quantiles of temperature distributions are estimated as follows (definition 7 in Hyndman and Fan, 1996), using O as an example. Let $\{O_1, O_2, \dots, O_n\}$ be samples from the distribution of O , $O_{(1)} \leq \dots \leq O_{(n)}$ be the order statistics, and O_p be the p quantile of O such that $F_O(O_p) = p = Pr(O \leq O_p)$. Then O_p is estimated by the weighted average of consecutive order statistics

$$\hat{O}_p = (1 - \omega)O_{(\lfloor np + (1-p) \rfloor)} + \omega O_{(\lfloor np + (1-p) \rfloor + 1)},$$

where $\lfloor \cdot \rfloor$ represents the integer part, and the weight ω is given by

$$\omega = np + (1 - p) - \lfloor np + (1 - p) \rfloor.$$

Other quantile estimators are available (e.g. Hyndman and Fan, 1996), but given the large sample sizes of temperature distributions (daily data from 30 summers) considered in this thesis, the results should be insensitive to the choice of estimators.

In addition, two quantities proposed by Ferro et al. (2005), ‘location-adjusted quantile difference’ and ‘location and scale-adjusted quantile difference’, are considered. Take BC-L and BC-LS as examples. For BC-L, if the assumptions (5.1) and (5.2) hold, i.e. G and O are only different in location, the location-adjusted quantile difference,

$$G_p - [\alpha_G + (O_p - \alpha_O)], \quad (5.16)$$

is expected to be zero for all $p \in (0, 1)$. On the other hand, if the assumption (5.4) of BC-LS holds, i.e. G and O are only different in location and scale, the location and scale-adjusted quantile difference,

$$G_p - \left[\alpha_G + \beta_G \left(\frac{O_p - \alpha_O}{\beta_O} \right) \right], \quad (5.17)$$

is expected to be zero for all $p \in (0, 1)$. In (5.16) and (5.17), G_p and O_p are estimated using sample quantiles, while the location and scale parameters are estimated by the sample mean and standard deviation, as described in Section 5.3.1. Plots of these two quantities against p and maps of these quantities for selected p are then used to validate the calibration assumptions. The assumptions of the change factor approach can be validated similarly.

The sampling uncertainties of quantile differences are estimated using a bootstrap resampling approach. Following Ferro et al. (2005), a block resampling technique is applied to account for serial dependence within each summer. With this technique, blocks of temperature values from each summer are resampled with replacement. Pointwise confidence intervals of quantile differences are then constructed from a large number (999 for this thesis) of resamples. Details on this approach are described in Chapter 8 of Davison and Hinkley (1997).

5.3.3 Estimation of return levels of extreme temperatures

The magnitudes of temperature extremes calibrated by different approaches are reported in terms of the estimated *return levels*. A return level refers to the quantile expected to be exceeded with a probability $(1 - p)$, which is expressed in terms of the return period $1/(1 - p)$. For example, for a summer with a length of 120 days, the temperature return level with a return period of 2 summers, or the ‘2-summer return level’, is the temperature level expected to be exceeded, on average, once every 240 summer days. ‘Moderate’ extreme return level, such as the level with return period of 1 summer, can be estimated using the sample quantile. For the estimation of rarer extremes with longer return periods (e.g. 5 or 10 summers), fitting a parametric model is more appropriate (Folland and Anderson, 2002). In this case, return levels are estimated by fitting generalized Pareto (GP) distribution to temperature excesses above a threshold. Consider again, O as an example. Let $V_O = O - u_O$ be the excesses of O above a sufficiently high threshold u_O , then the GP distribution is given by

$$Pr(V_O \leq v_O | V_O > 0) = 1 - \left(1 + \frac{\xi_O v_O}{\tilde{\sigma}_O}\right)^{-\frac{1}{\xi_O}},$$

where $\tilde{\sigma}_O$ and ξ_O are scale and shape parameters respectively, and $v_O > 0$ and $(1 + \xi_O v_O / \tilde{\sigma}_O) > 0$. In this thesis, a constant threshold of sample 0.95 quantile of O is used for u_O . Then for a given return period r_m (in days), the temperature return level T_m are estimated by

$$\hat{T}_m = \begin{cases} u_O + \frac{\hat{\sigma}_O}{\hat{\xi}_O} \left[(0.05r_m)^{\hat{\xi}_O} - 1 \right] & \text{for } \hat{\xi}_O \neq 0 \\ u_O + \hat{\sigma}_O \log(0.05r_m) & \text{for } \hat{\xi}_O = 0, \end{cases}$$

where $\hat{\sigma}_O$ and $\hat{\xi}_O$ are estimates of $\tilde{\sigma}_O$ and ξ_O respectively. Other temperature variables are treated similarly. Details of the GP distribution and the estimation of return levels are given in Appendix C.

5.4 Calibration results for London

In this and the following section, the calibration for the standard version of HadRM3 is first presented. This is followed by results involving the other members of the perturbed physics ensemble.

5.4.1 Comparison of temperature distributions

The quantile-quantile plot of G against O , shown by the grey dots in Fig. 5.3(a), can be used to evaluate the bias of the present-day HadRM3 simulated temperatures. This plot follows roughly a straight line close to the line of equal values up to around $O = 23$ °C, above which a curvature appears. This indicates that compared to observations, the distribution of modelled temperatures has a heavier tail at high temperatures. This feature is also apparent when comparing the boxplots of O and G [the first two boxes in Fig. 3.6(a)]. The quantile difference plot [Fig. 5.3(b)] shows that the bias of G is negative (cold) and small (about 0.5 °C) at the intermediate range of temperatures, but a positive (warm) bias of more than 3 °C can be observed for the hot extreme. This can be seen more clearly in Fig. 5.3(c) where the sample quantile difference is plotted against the return period. As will be discussed in Section 5.4.2, this bias continues to increase for rarer extremes, i.e. with longer return periods. Both the location-adjusted quantile difference and location

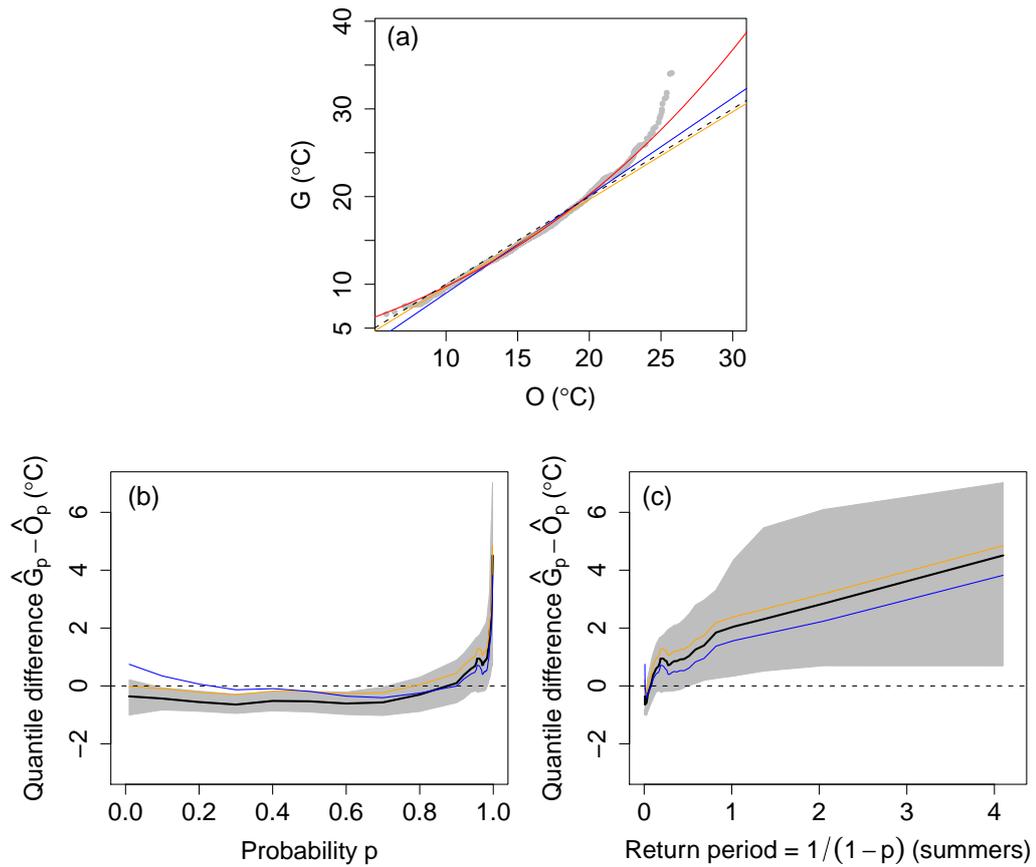


Figure 5.3: (a) Quantile-quantile plot (grey dots) of G of HadRM3 standard version against O for London. The dashed lines indicate equal values of G and O . The mapping of G' values (or G for bias correction of the present-day values) onto O' values by the three bias correction methods are shown by coloured lines: orange for BC-L (5.3), blue for BC-LS (5.5), red for BC-LSB (5.9). (b) Sample quantile difference between G and O against probability for London. The 90% pointwise bootstrap confidence intervals are indicated by the grey shaded bands. The location-adjusted sample quantile difference (5.16) and location and scale-adjusted sample quantile difference (5.17) are superimposed by orange and blue lines respectively. The corresponding confidence intervals are not shown for clarity. (c) As in panel (b), but with the probability expressed in terms of return periods, so that the extreme quantile differences can be more easily observed.

and scale-adjusted quantile difference, shown by orange and blue lines respectively, are similar in magnitude to the quantile difference. These suggest that the warm bias of G at

the tail cannot be explained by its bias in location and scale alone and is likely to be related to the difference between the shapes of G and O . The discussion of possible causes of such a difference is postponed to Section 5.6.1 when G and O are compared for the whole HadRM3 domain.

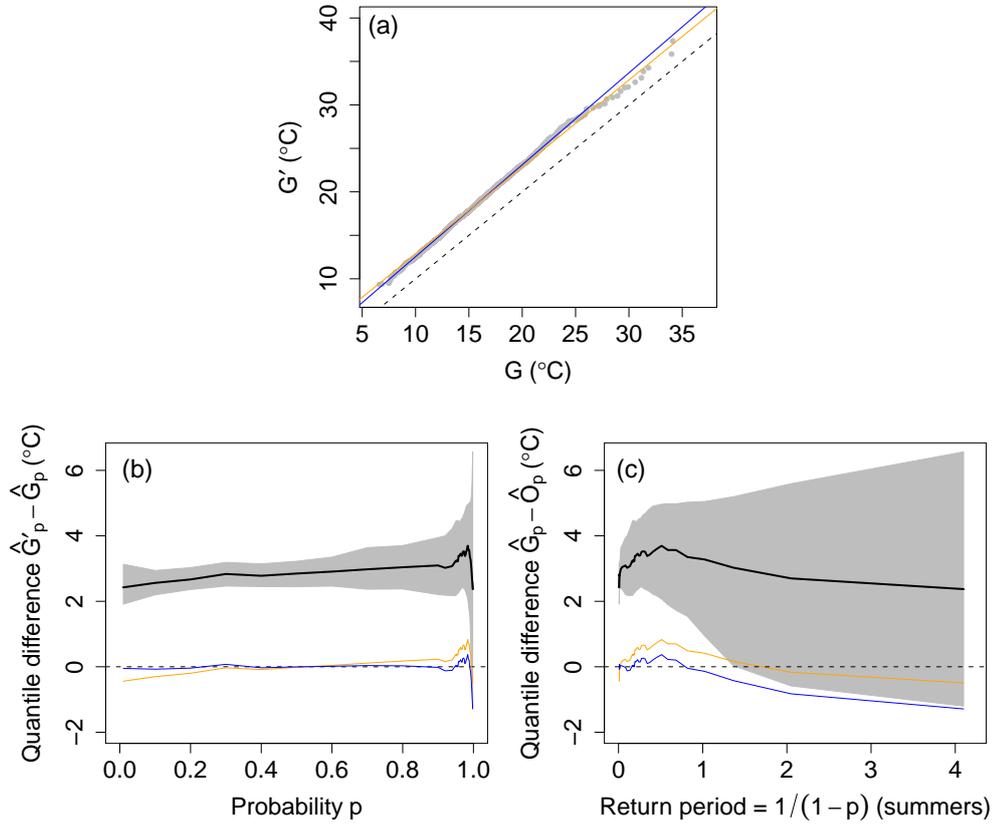


Figure 5.4: (a) Quantile-quantile plot of G' (2040 to 2069) against G for London. (b) and (c) Sample quantile difference between G' and G (2040 to 2069) against probability and return period for London. Refer to the caption of Fig. 5.3 for further explanation of these plots.

The change factor method is based on the relationship between HadRM3 temperatures for the present-day period and the specific future time-slice (G and G'). For the sake of brevity, only the quantile-quantile plot and quantile difference plot involving G' for the 2040 to 2069 time-slice are shown in Fig. 5.4. The plots involving the other two time-slices display similar features. The quantile-quantile plot [Fig. 5.4(a)] follows a straight line roughly parallel to the line of equal values, apart from some minor departure at the hot

extreme. This suggests that the distributions of G and G' are mainly different in location. This is confirmed by Figs 5.4(b) and (c) which show that the estimated quantile difference ($\hat{G}'_p - \hat{G}_p$) is around 2.5 °C across all values of p , i.e. a warming of about 2.5 °C of modelled temperatures between the two time-slices. Both the location-adjusted quantile difference and location and scale-adjusted quantile difference are close to zero for most values of p .

5.4.2 Extreme temperatures for standard HadRM3 simulation

The return level plots in Fig. 5.5 show the effects of various calibration methods on extreme temperatures. The return curves are based on GP distribution fits to excesses of different temperature variables above thresholds for different time periods, as described in Section 5.3.3. For the evaluation of the model fits, the empirical return level estimates are overlaid by the grey symbols. For certain variables, for example G and G' for the 2070 to 2099 time-slice in Fig. 5.5(f), the model fitted return curves (black solid and dashed lines) have some departure from the empirical estimates. Choosing a threshold higher than the 0.95 sample quantile appears to produce better fits in such cases. Such results are not presented here so that the results shown in Fig. 5.5 are consistent with that to be presented in Section 5.6. Nevertheless, when referring to temperature extremes, the discussion below focuses on the estimated 10-summer return level, which is not very sensitive to the choice of threshold.

For the present-day, the 10-summer return level of HadRM3 temperatures (G) has a warm bias of about 6 °C compared to observations (O) [Fig. 5.5(f)]. Since there are little differences in location and scale between O and G , as shown by the estimates of sample mean and standard deviation in Table 5.2, such bias is not much reduced after applying BC-L or BC-LS [Figs 5.5(a) and (b)]. The BC-LSB method, however, gives considerable improvement [Fig. 5.5(c)]. The bias of the 10-summer level is reduced to less than 2 °C. The mapping of G values onto O' values by BC-LSB is shown by the red line in Fig. 5.3(a). This is closer to the relationship between the sample quantiles of O and G compared to the mapping by BC-L and BC-LS which are shown by orange and blue lines respectively, even though some discrepancy is still apparent at the hot extreme. This suggests that it is

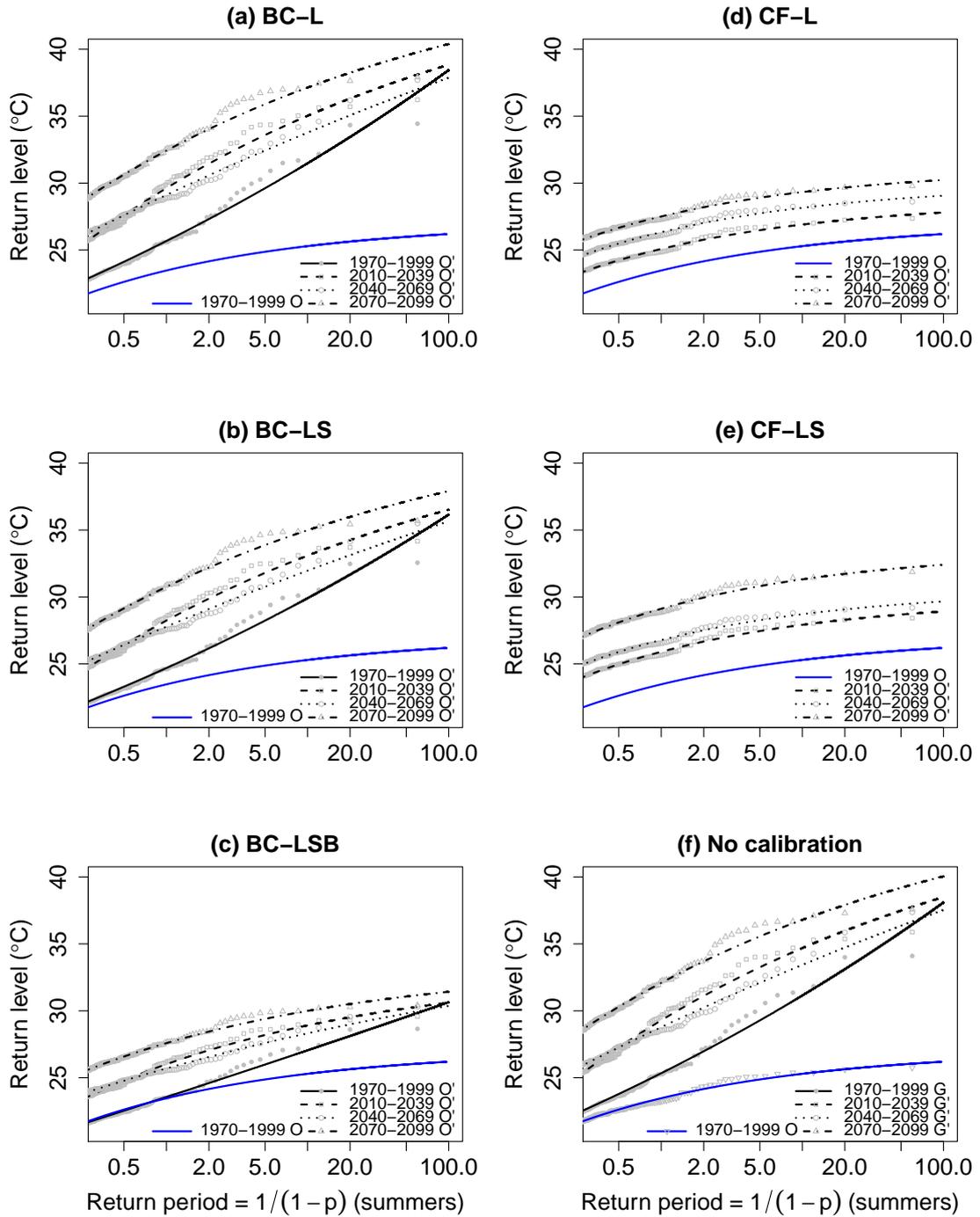


Figure 5.5: Return level plots of different temperature variables for various time-slices (as labelled) for London. Panels (b) through (f) show the results using different calibration methods. On each panel the return curve for O is superimposed in blue. The empirical estimates of return levels are shown by grey symbols for the evaluation of GP distribution fits.

Variable	Time period	Mean (°C)	s.d. (°C)	10-summer level (°C)
(a) No calibration				
O	1970-1999	15.9	3.01	25.3 [25.0,25.8]
G	1970-1999	15.6	3.35	31.1 [29.6,33.8]
G'	2010-2039	17.2	3.70	34.7 [33.5,36.9]
G'	2040-2069	18.5	3.54	33.4 [32.4,35.2]
G'	2070-2099	19.6	4.03	36.8 [35.7,38.7]
(b) BC-LS				
O'_b	1970-1999	15.9	3.01	29.9 [28.5,32.3]
O'_b	2010-2039	17.4	3.33	33.1 [32.0,35.1]
O'_b	2040-2069	18.5	3.18	32.0 [31.0,33.6]
O'_b	2070-2099	19.6	3.63	35.0 [34.0,36.7]
(c) BC-LSB				
O'_b	1970-1999	15.9	3.01	27.0 [26.2,28.6]
O'_b	2010-2039	17.4	3.08	28.9 [28.3,30.0]
O'_b	2040-2069	18.4	2.84	28.3 [27.8,29.2]
O'_b	2070-2099	19.3	3.06	30.0 [29.5,30.9]
(d) CF-LS				
O'_c	2010-2039	17.6	3.33	27.9 [27.5,28.5]
O'_c	2040-2069	18.8	3.18	28.7 [28.3,29.3]
O'_c	2070-2099	20.0	3.63	31.3 [30.9,32.0]

Table 5.2: Sample mean, standard deviation (s.d.) and estimated 10-summer return level (from GP distribution fit) of summer temperatures for London (a) with no calibration and (b), (c) and (d) calibrated using three different methods. The numbers in square brackets show the 90% confidence intervals of 10-summer return level, which are estimated by the delta method as described in Chapter 4 in Coles (2001).

more suitable to apply bias correction in location and scale on transformed temperatures because the Box-Cox transformation caters for the different shapes between O and G to a certain extent. Therefore, among the three bias correction methods, only the results obtained by BC-LSB are considered to be appropriate.

If no calibration is applied, the 10-summer return level projected by HadRM3 will be 36.8 °C in the period 2070 to 2099, an increase of 11.5 °C relative to O in the period 1970 to 1999 (Table 5.2). The corresponding projected increase in the mean summer temperature between the two periods is only 3.7 °C. After applying BC-LSB, however, the 10-summer level in 2070 to 2099 is 30.0 °C, which is only a 4.7 °C higher than that of O . Such a large discrepancy has important implications on heat-related mortality projections. For example, using mortality model TS developed for London in Chapter 4, if the daily mean temperature was 36.8 °C on a day in mid-summer 2003, the predicted number of elderly deaths would be 481 (95% confidence interval: [359,602]). In contrast, if the daily mean temperature was 30.0 °C, the predicted number of deaths would be 247 (95% confidence interval: [217,278]), which is almost 50% lower.

For the change factor approach, the difference between the estimated extreme return levels of temperatures calibrated by the two methods (CF-L and CF-LS) is around 1 to 2 °C [compare Figs 5.5(d) and (e)]. Since the transfer function of the change factor approach is identity for no change in time, the reliability of the two methods cannot be compared by applying them on G as has been done for the bias correction approach. In this case, the quantile difference plots [Figs 5.4(b) and (c)] are useful diagnostics as they can be used to assess the validity of the assumptions. On these plots, the location and scale-adjusted quantile difference is slightly closer to zero than the location-adjusted quantile difference for most values of p , therefore the assumptions of CF-LS (5.11) are considered more plausible.

The 10-summer level of temperatures calibrated by CF-LS is 31.3 °C in 2070 to 2099 [Table 5.2(d)], 6.0 °C higher than that of O . Comparing the results of BC-LSB and CF-LS methods in Table 5.2(c) and (d), the differences in the estimated 10-summer levels for the three future periods range from 0.4 to 1.3 °C. These are not significant considering the sampling uncertainties indicated by the 90% confidence intervals.

5.4.3 Extreme temperatures for ensemble members

Temperature projections from each of the other 10 members of the PPE are calibrated individually in the same way as the standard member. Quantile-quantile plots and quantile difference plots involving these members are not shown here for conciseness. However, as the boxplot of O and G for all PPE members [Fig. 3.6(a)] shows, the simulated temperatures of all PPE members have heavier tails at high temperatures compared to observations. These differences cannot be reasonably explained by the differences in the location and scale between O and G . As such, BC-LSB is preferred among the three bias correction methods. Meanwhile, the changes in the distribution of projected temperatures for all PPE members of HadRM3 with time mainly involve changes in location and scale (not shown). The assumptions of the CF-LS method are therefore satisfied.

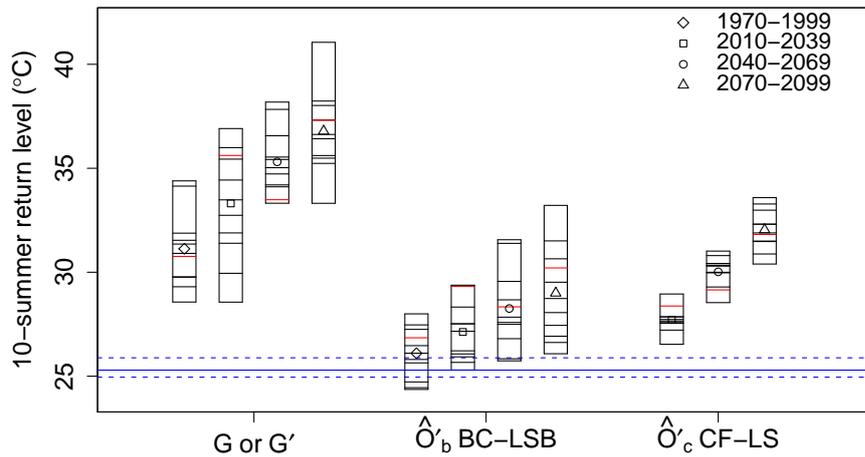


Figure 5.6: Estimates of 10-summer return level of uncalibrated HadRM3 temperature projections (G and G'), projections calibrated by BC-LSB (\hat{O}'_b) and projections calibrated by CF-LS (\hat{O}'_c) in London for each member of PPE. The symbols in each box represent the mean of the return level estimates of the PPE members for each 30-year period as labelled. The black horizontal lines within each box show the return level estimates of each PPE member, with the estimate for the standard run shown in red. The blue solid line represents the 10-summer return level estimates for O in 1970 to 1999, and the dashed lines represent its 90% confidence interval.

The estimated 10-summer return levels of calibrated temperatures using BC-LSB and

CF-LS for the PPE members are shown in Fig. 5.6. The 10-summer levels of uncalibrated temperatures (G for present-day and G' for the three future periods) are also shown. The length of each box shows the spread of 10-summer level estimates among the 11 PPE members in the specified 30-year period, while estimates of individual PPE members are marked by horizontal lines within each box. The mean of the 11 estimates ('ensemble mean') are shown by different symbols for each period. The ensemble mean bias in the 10-summer level of G is about 6 °C compared to O . After being calibrated by BC-LSB, this bias is reduced to about 1 °C. For the future, the ensemble mean 10-summer level of \hat{O}'_b using BC-LSB in the period 2070 to 2099 is 3.7 °C higher relative to O . However, the estimates of increase among the PPE members range from 0.8 to 7.9 °C. This reflects that there are considerable uncertainties, which arise from the uncertainty in climate model parameters, in the projections of future changes in extreme temperatures calibrated using this technique. In contrast, the ensemble spread for the 10-summer level of \hat{O}'_c using CF-LS is smaller. For the period 2070 to 2099, this ranges from 5.1 to 8.2 °C (with an ensemble mean of 6.7 °C) higher than the 10-summer level of O . This spread is smaller than the spread of G' for the same period. It appears that the climate model uncertainty in extreme temperature projections can be reduced by applying the change factor calibration approach.

5.5 Calibration results for Budapest

We now consider the calibration of HadRM3 temperatures for Budapest. This is of interest not only because Budapest is the other focus city of this thesis, but also because in terms of the distributional properties, the nature of HadRM3 temperature biases is different between Budapest and London.

5.5.1 Comparison of temperature distributions

The quantile-quantile plot of G against O for Budapest [Fig. 5.7(a)] follows roughly a straight line which is steeper than the line of equal values. This suggests that the variance

(scale) of the distribution of G is larger than that of O , and thus explains the increasing quantile difference with p [Fig. 5.7(b)] which is more than $7\text{ }^\circ\text{C}$ at high temperatures [Fig. 5.7(c)]. The location and scale-adjusted quantile difference is less than $0.5\text{ }^\circ\text{C}$ for all p except at the low temperature extreme. The bias in the distribution of G can therefore be well explained by biases in both location and scale.

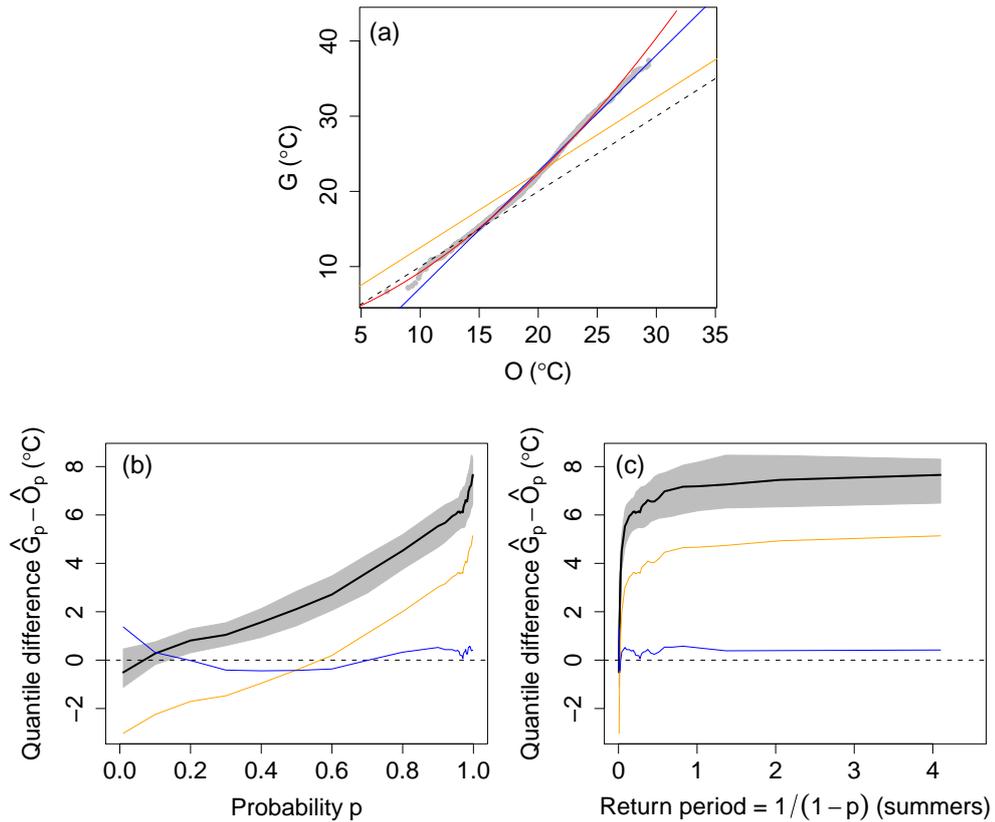


Figure 5.7: As in Fig. 5.3 but for Budapest.

As for the distributions of G' (for the time-slice 2040 to 2069) and G for Budapest, the quantile-quantile plot in Fig. 5.8(a) does not indicate differences between their shapes. From Figs 5.8(b) and (c), the warming of modelled temperatures between the two periods is about 3 to 4 $^\circ\text{C}$. Similar to London, the location-adjusted quantile difference is reasonably close to zero for all p , suggesting that the two distributions are mainly different in location. However, as the location and scale-adjusted quantile difference is even closer to zero, G and G' may have a small difference in their scale.

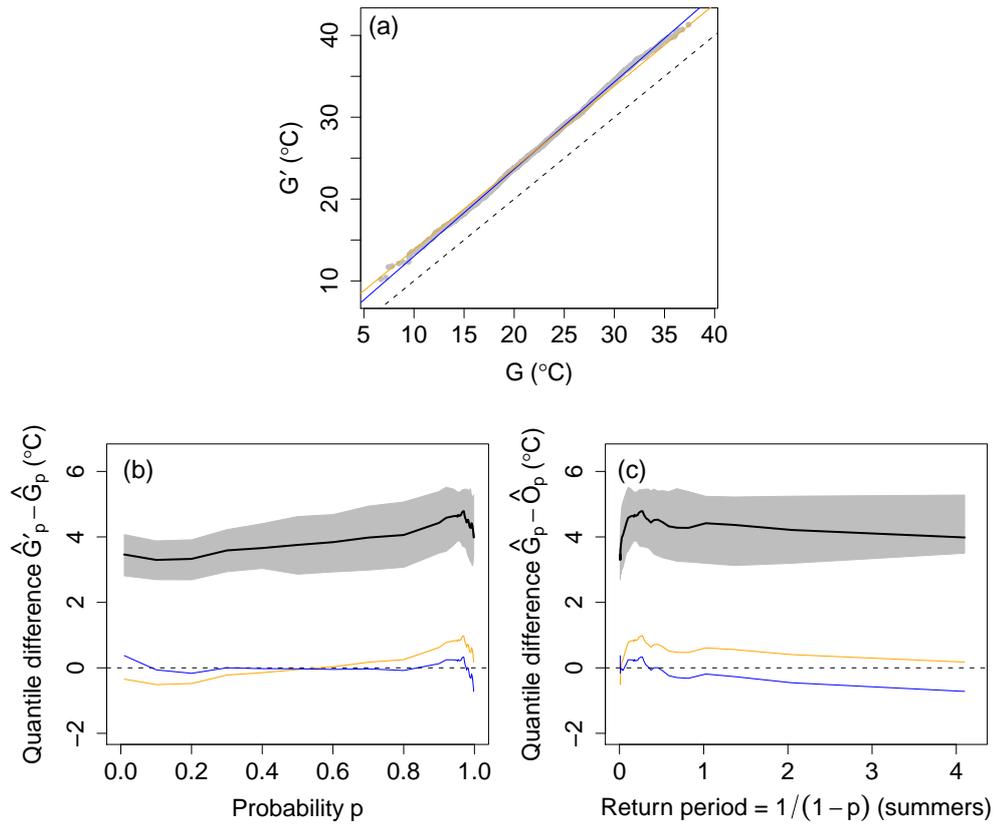


Figure 5.8: As in Fig. 5.4 but for Budapest.

5.5.2 Extreme temperatures for standard HadRM3 simulation

The simulated temperatures of HadRM3 have warm biases of about 8 °C at the extreme compared to observations [Fig. 5.9(f)]. Such biases are removed after correcting the biases in both the location and scale of G using BC-LS [Fig. 5.9(b)]. In addition, comparing the blue and red lines in Fig. 5.7(a), which show the mapping of G' values onto O' values for the calibration without (BC-LS) and with Box-Cox transformation (BC-LSB) respectively, the former is slightly closer to the relationship between the sample quantiles of O and G . These two pieces of evidence suggest that the use of BC-LSB is not necessary in this case.

As in the case of London, calibration using bias correction considerably lowers the future projected changes in extreme temperatures by HadRM3. From Table 5.3, the model projected (G') 10-summer level in the period 2070 to 2099 is 14.0 °C higher compared to the observed (O) 10-summer level in the period 1970 to 1999. In contrast, the 10-summer

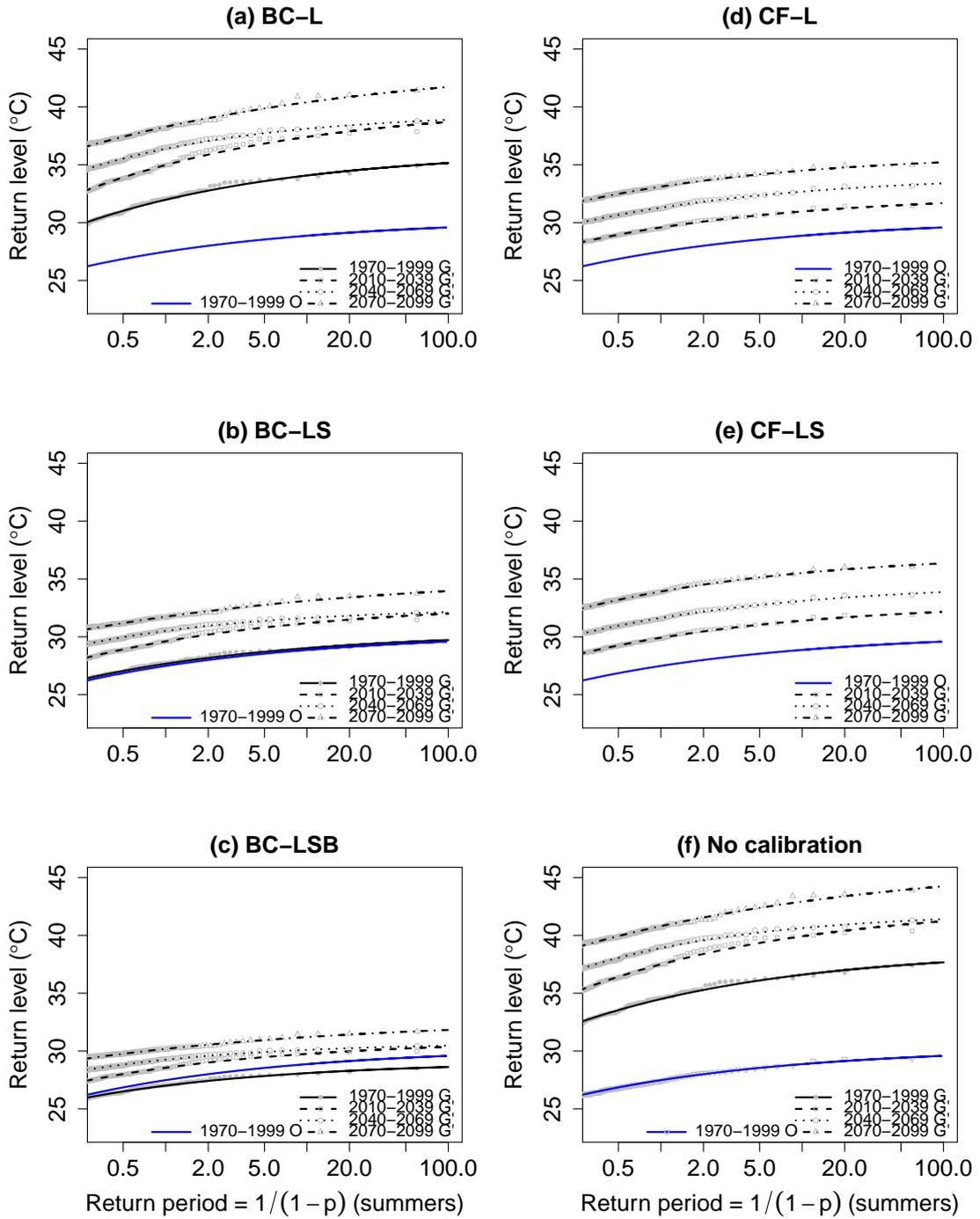


Figure 5.9: As in Fig. 5.5 but for Budapest.

level of bias-corrected temperatures \hat{O}_t^b using BC-LS in the 2070 to 2099 period is only 5.0 °C higher than that of O .

The two change factor methods give similar results for future changes in extreme tem-

Variable	Time period	Mean (°C)	s.d. (°C)	10-summer level (°C)
(a) No calibration				
O	1970-1999	19.8	3.52	28.9 [28.6,29.3]
G	1970-1999	22.3	5.45	36.7 [36.2,37.4]
G'	2010-2039	24.4	5.81	39.9 [39.5,40.9]
G'	2040-2069	26.1	5.80	40.6 [40.1,41.2]
G'	2070-2099	28.0	6.30	42.9 [42.5,43.6]
(b) BC-LS				
O'_b	1970-1999	19.8	3.52	29.0 [28.7,29.4]
O'_b	2010-2039	21.2	3.75	31.2 [30.9,31.7]
O'_b	2040-2069	22.3	3.75	31.6 [31.3,32.0]
O'_b	2070-2099	23.4	4.07	33.1 [32.8,33.5]
(c) BC-LSB				
O'_b	1970-1999	19.8	3.52	28.1 [27.9,28.5]
O'_b	2010-2039	21.1	3.59	29.8 [29.5,30.2]
O'_b	2040-2069	22.1	3.47	30.1 [29.9,30.3]
O'_b	2070-2099	23.2	3.68	31.2 [31.0,31.5]
(d) CF-LS				
O'_c	2010-2039	21.7	3.75	31.4 [31.1,31.8]
O'_c	2040-2069	23.4	3.75	33.1 [32.5,33.5]
O'_c	2070-2099	25.0	4.07	35.5 [35.2,36.0]

Table 5.3: Same as Table 5.2 but for Budapest.

peratures [Figs 5.9(d) and (e)]. Comparing the location-adjusted quantile difference and location and scale-quantile difference in Fig. 5.8(b), the latter is closer to zero for most values of p . Therefore, the CF-LS method is preferred.

From Table 5.3, the mean of calibrated temperatures using CF-LS (\hat{O}'_c) are 0.5 to 1.6 °C higher than that using BC-LS (\hat{O}'_b). These are consistent with the estimated difference between $\mathbb{E}(O'_b)$ and $\mathbb{E}(O'_c)$ using (5.14) (calculations not shown). In addition, the standard

deviation of \hat{O}'_b and \hat{O}'_c are the same, which is another result expected from (5.15). The difference between the mean of \hat{O}'_b and \hat{O}'_c can, at least partly, explain the 0.2 to 2.4 °C difference in their 10-summer return levels.

5.5.3 Extreme temperatures for ensemble members

For the other 10 members of the PPE, there is little difference between the shapes of O and G and between the shapes of G and G' (not shown). BC-LS and CF-LS are therefore the most suitable methods under each calibration approach. All members of the PPE have large warm biases in the present-day simulations of 10-summer return levels, ranging from 7.7 °C to 11.3 °C (Fig. 5.10). These are well corrected by the BC-LS method, with the magnitudes of biases of \hat{O}'_b for the present-day being less than 2 °C for all PPE members. Compared to London, the ensemble spread in projected future warming in the extremes calibrated by the bias correction approach is slightly smaller. The 10-summer level of \hat{O}'_b in the period 2070 to 2099 is 1.0 to 6.0 °C (ensemble mean: 3.6 °C) higher than O in the present-day. The corresponding estimates using the CF-LS are generally higher, ranging from 4.6 to 8.3 °C (ensemble mean: 6.3 °C).

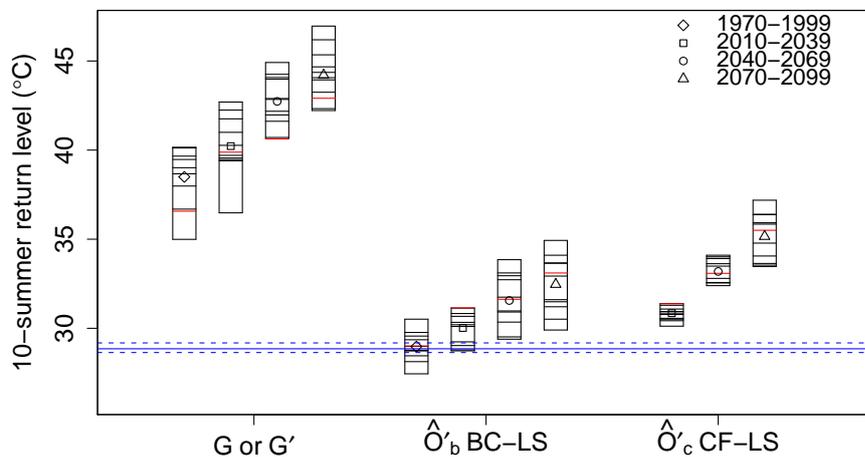


Figure 5.10: Similar to Fig. 5.6 but for Budapest. Here the estimated 10-summer levels of uncalibrated HadRM3 temperature projections (G and G'), projections calibrated by BC-LS (\hat{O}'_b) and projections calibrated by CF-LS (\hat{O}'_c) are shown.

5.6 Spatially extended calibration over Europe

In this section, we consider the temperature calibration for the whole HadRM3 domain over Europe. Although the results to be presented in this section are not used to project mortalities for London and Budapest in Chapter 6, these are included in this thesis for two reasons. First, as mentioned in Section 1.2, such results can give an indication of how the risks of heat-related mortality will change over Europe. Second, as we will see later, the results further demonstrate the importance of the choice of calibration approach in mortality projections.

5.6.1 Comparison of present-day observed and simulated temperatures

We first compare the spatial maps of three summary statistics for location, scale and shape of O and G . Figures 5.11(a) and (b) show the sample mean of O and G respectively. The main spatial patterns of observed mean temperatures are generally well simulated by HadRM3, including the north-south gradient and the lower temperatures over elevated terrain (e.g. the Alps). However, the RCM has a warm bias of around 2 to 4 °C over Italy and southern Europe, and has a small cold bias over parts of Scandinavia [Fig. 5.11(c)]. The variance of the two distributions are compared by the standard deviation [Figs 5.11(d) and (e)]. HadRM3 overestimates the present-day variance over most parts of continental Europe, especially in the south where the standard deviation of G is more than 50% greater than that of O [Fig. 5.11(f)]. The shapes of O and G are compared by their sample moment skewness [Figs 5.11(g) and (h)]. The observed temperatures are positively skewed (longer tail at the high temperature end) over northern France, the Netherlands and Germany, but are negatively skewed (shorter tail at the high temperature end) in Spain, Italy and over eastern Europe [Fig. 5.11(g)]. The map of sample skewness of modelled temperatures [Fig. 5.11(h)] shows the same general spatial patterns, except for the negative skewness over eastern Europe. However, the modelled temperatures are more positively skewed compared to observations in England and near the northern coast of France and Spain

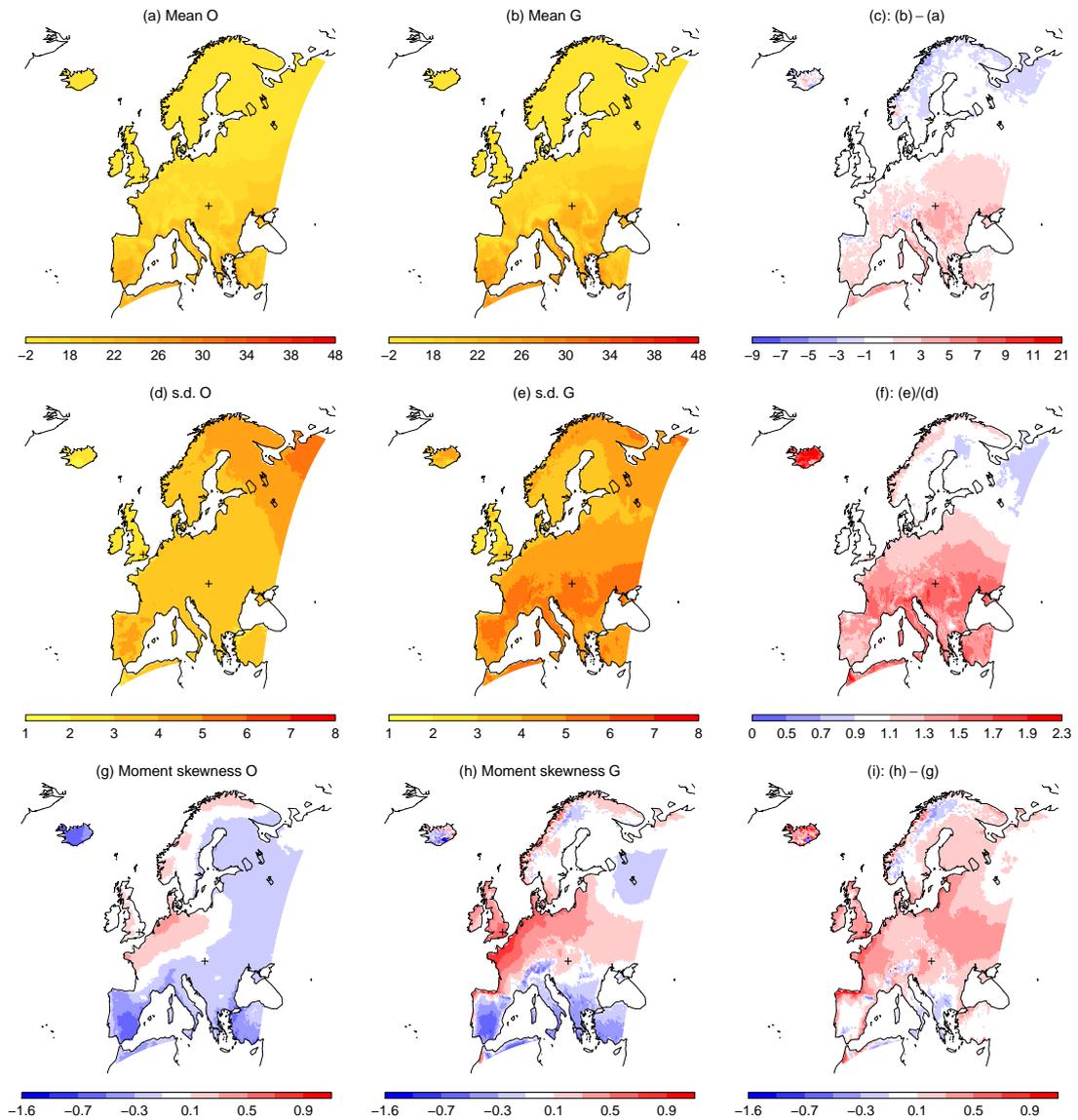


Figure 5.11: Maps showing the following sample estimates of O and G for the present-day: (a) and (b) mean (in $^{\circ}\text{C}$); (d) and (e) standard deviation (in $^{\circ}\text{C}$); (g) and (h) moment skewness. The difference in the sample mean [(b) - (a); in $^{\circ}\text{C}$], the ratio of standard deviation (e)/(d) and the difference in moment skewness [(h) - (g)] are shown in panels (c), (f) and (i) respectively. On these and the other maps in this chapter, the (spatial) locations of London and Budapest are marked by '+' for reference.

[Fig. 5.11(i)].

The magnitude of bias of G for the present-day increases towards the hot extreme. This can be seen by comparing the maps showing the extreme quantile differences [Figs 5.12(a) and (b)] with the map showing the bias in the sample mean [Fig. 5.11(c)]. The warm bias for the 0.99 quantile (corresponding to a return period of approximately 1 summer) is greater than $5\text{ }^{\circ}\text{C}$ over many parts of continental Europe, and the bias for the 10-summer return level [Fig. 5.12(e)] even exceeds $7\text{ }^{\circ}\text{C}$. A warm bias at the hot extreme also appears over northern France and southern England.

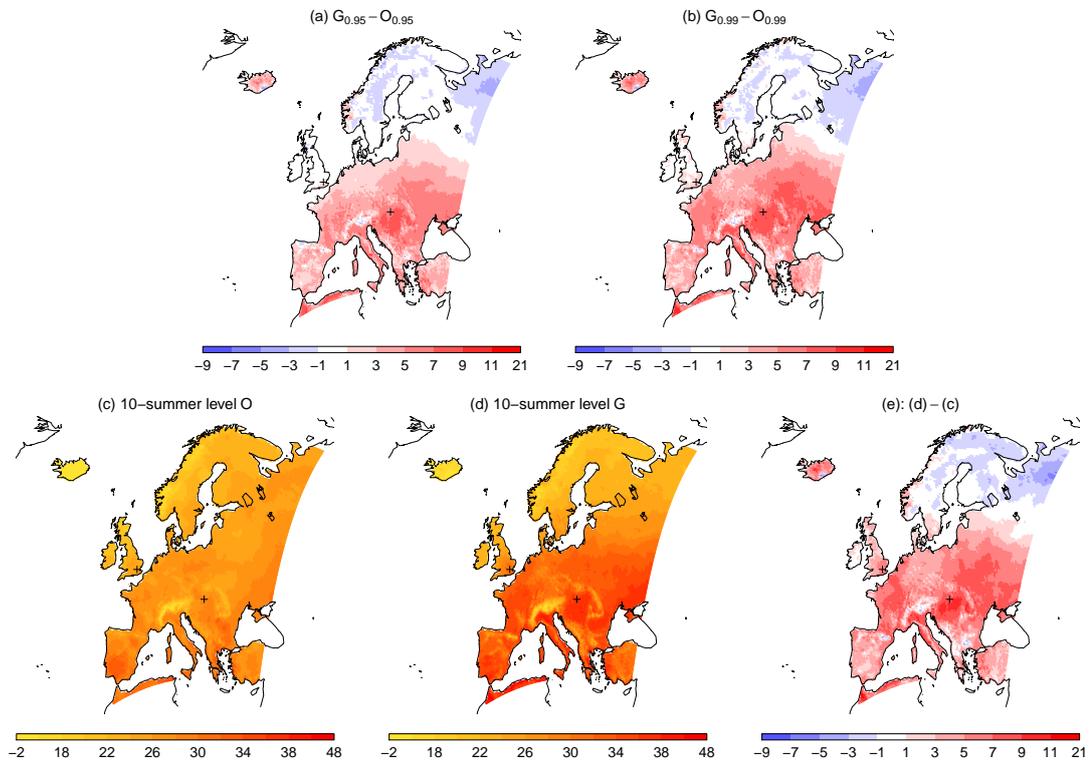


Figure 5.12: Maps comparing extreme G and O . (a) and (b): differences in sample 0.95 and 0.99 quantile respectively between G and O (in $^{\circ}\text{C}$). The 10-summer return level (in $^{\circ}\text{C}$) estimated by GP distribution fits for O and G are shown in (c) and (d) respectively. Their difference [(d) - (c) in $^{\circ}\text{C}$] is shown in (e).

After adjusting the difference in location between G and O , the 0.99 quantile difference is still generally greater than $3\text{ }^{\circ}\text{C}$ over central Europe [Fig. 5.13(a)], while the location and scale-adjusted quantile difference is less than $1\text{ }^{\circ}\text{C}$ [Fig. 5.13(b)]. A similar pattern is also

observed for the location and scale-adjusted 0.95 quantile difference (not shown). These suggest that in this region, the biases in the extreme can be explained by the model biases in both the location and scale of the temperature distributions. However, the location and scale-adjusted 0.99 quantile difference remains greater than 1 °C in northern France, southern England and the northeastern part of Europe. These are the regions where G is more positively skewed compared to O as discussed earlier. The model biases at the extreme for these regions are therefore likely to be related to differences between the shapes of the G and O distributions.

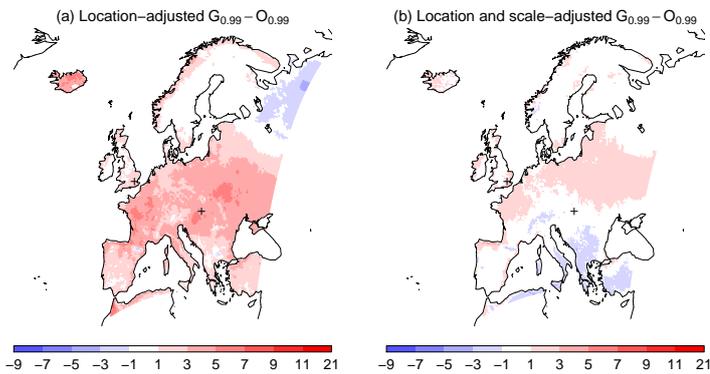


Figure 5.13: (a) Location-adjusted (5.16) and (b) location and scale-adjusted 0.99 quantile difference (5.17) (in °C) between G and O .

5.6.2 Causes of HadRM3 biases for extreme temperatures

The causes of the model biases in simulating present-day temperatures are explored by considering the present-day simulations of HadRM3 forced by ERA-40 re-analysis at the boundaries (T^a ; see Section 3.4.1). The sample mean, standard deviation, 0.99 quantile and moment skewness of T^a and O are compared in Fig. 5.14. The spatial patterns shown on these maps are generally consistent with the comparison of these sample statistics between G and O in Figs 5.11(c), (f), (i) and 5.12(b). HadRM3 driven by ERA-40 re-analysis overestimates the mean and standard deviation of temperatures over central and southern Europe [Figs 5.14(a) and (b)]. The bias in the 0.99 quantile is greater in magnitude compared to the mean [Fig. 5.14(c)]. However, the magnitudes of biases in

these three statistics for this model run are generally smaller than that for HadRM3 forced by HadCM3. Similar to G , the distribution of T^a is more positively skewed compared to that of O , especially in the northern coastal regions of Spain and France [Fig. 5.14(d)]. If the boundary forcing provided by re-analysis data is assumed to be ‘perfect’, the above results suggest that the HadRM3 biases in temperature simulations are related, at least partly, to the RCM formulation.

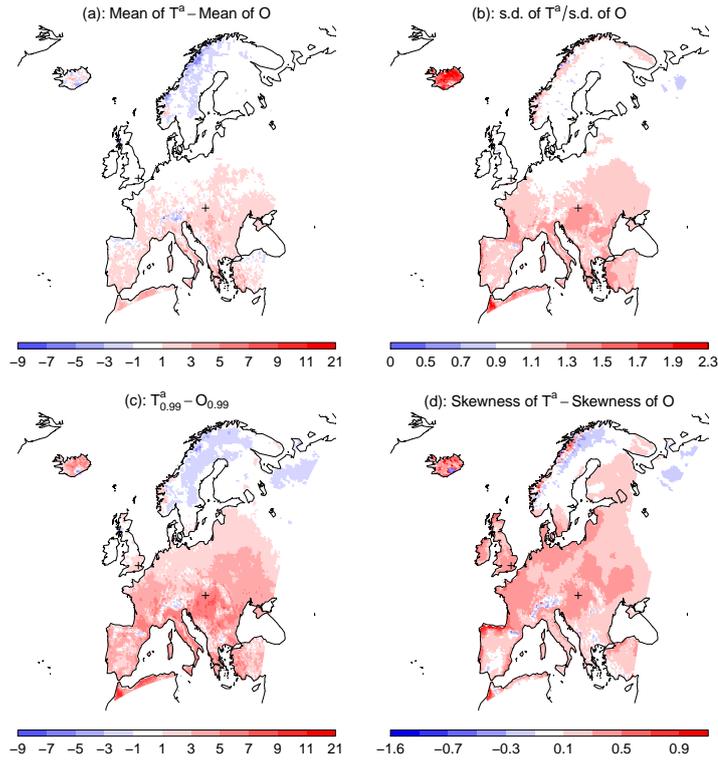


Figure 5.14: (a) Difference in sample mean (in $^{\circ}\text{C}$), (b) ratio of standard deviation, (c) difference in sample 0.99 quantile (in $^{\circ}\text{C}$) and (d) difference between moment skewness between T^a and O .

The problem of increasing model summer temperature biases at high temperatures has been observed in other RCMs. Christensen et al. (2008) compared monthly mean temperatures simulated from 13 different RCMs forced by ERA-40 re-analysis data with E-OBS observed monthly mean temperatures. Model biases were found to increase non-linearly with increasing observed temperatures for the majority of the RCMs. Such a relationship was found to be described by a second-order polynomial fit. As part of the

PRUDENCE project, Kjellström et al. (2007) compared summer (June to August) daily maximum temperatures simulated by 10 RCMs forced by the same boundary conditions with station observations. The biases at different temperature quantiles were examined. It was also found that the magnitudes of warm biases increase with temperatures at the hot extreme. Among the RCMs under consideration, the magnitude of bias was the greatest in HadRM3H, which is a similar RCM to the one used in this thesis but with a horizontal resolution of 50 km. Moberg and Jones (2004) found that large biases at the extreme for summer temperatures in the Hadley Centre RCM were associated with excessive soil drying in the model, especially over regions south of 45° N. Since drier soil limits evaporative cooling, the modelled temperatures increase too rapidly. The problem is made worse when a positive feedback of soil drying (refer to Section 2.2.2) develops in the model. Anders and Rockel (2009) suggested that excessive soil drying over southeastern Europe in RCMs is related to an inaccurate representation of soil type distribution in the model.

5.6.3 Comparison of HadRM3 temperatures in different time periods

The future changes in mean temperatures projected by HadRM3 are shown in Fig. 5.15. Relative to the simulated mean temperatures for the present-day G , the increase in mean temperatures is rather uniform over the model domain. Mean temperatures are projected to increase by about 2 °C in most parts of Europe in the period 2010 to 2039 [Fig. 5.15(d)]. Towards the end of the century, the increase in mean temperatures is projected to be larger in southern and eastern Europe (about 6 °C) and smaller elsewhere (about 4 °C) [Fig. 5.15(f)].

There is a small increase in the projected temperature variance for the future. For the first two future 30-year periods, the change in standard deviation of G' relative to G is less than 10% over most parts of Europe [Figs 5.16(d) and (e)]. By the period 2070 to 2099, the standard deviation of G' is around 20% greater than that of G generally over continental Europe and over the UK [Fig. 5.16(f)].

The comparison of sample moment skewness of G' with that of G is shown in Fig. 5.17. Unlike the changes in mean and variance which are rather spatially uniform, the distribu-

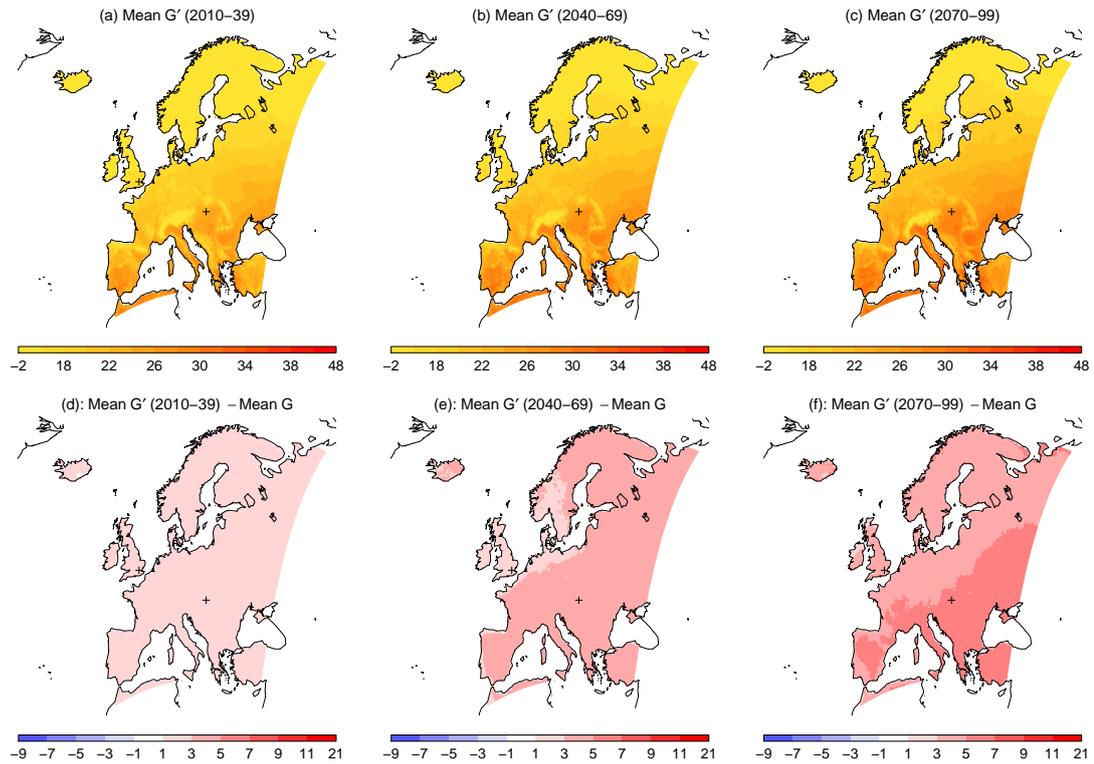


Figure 5.15: The sample mean of G' in the periods (a) 2010 to 2039, (b) 2040 to 2069 and (c) 2070 to 2099 (in $^{\circ}\text{C}$). The differences between the sample mean of G' for the three periods and that of G (in $^{\circ}\text{C}$) are shown in panels (d) through (f).

tion of HadRM3 temperatures is projected to become more positively skewed only in the northern part of Europe in the future. This is particularly the case over southern parts of Norway and Sweden, Denmark, the Netherlands, northern Germany, the UK and western Russia. The effects of changing skewness are clearly demonstrated by the maps showing the changes in the sample 0.99 quantile of HadRM3 temperatures (Fig. 5.18). The areas with the greatest increase in the 0.99 quantile are where the temperature distribution is projected to become more positively skewed. For example, by the period 2070 to 2099, the 0.99 quantile of modelled temperatures is projected to increase by about 8°C (relative to G) in Germany [Fig. 5.18(c)], where the increase in the mean is only about 4°C [Fig. 5.15(f)]. The map of location and scale-adjusted 0.99 quantile difference (not shown) suggests that this larger increase in the 0.99 quantile in these regions cannot be explained only by changes in location and scale of modelled temperature distributions.

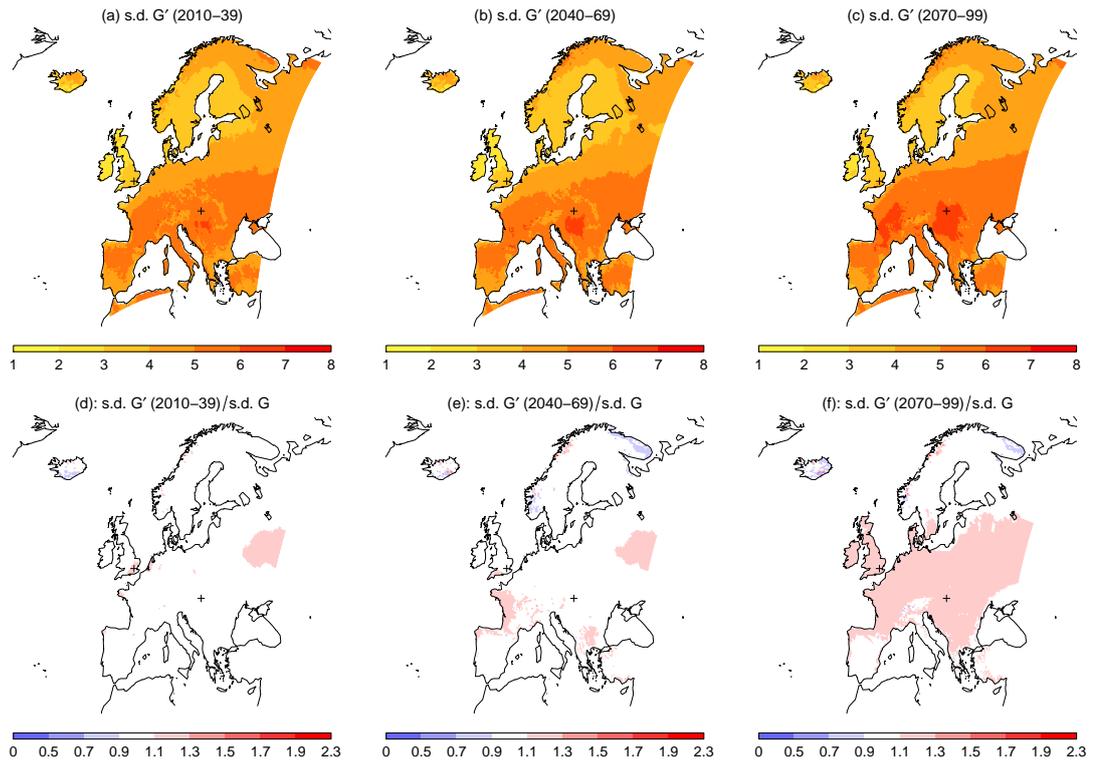


Figure 5.16: As in Fig. 5.15 but for standard deviation (in $^{\circ}\text{C}$). Panels (d) to (f) show the ratios of standard deviation between G' and G .

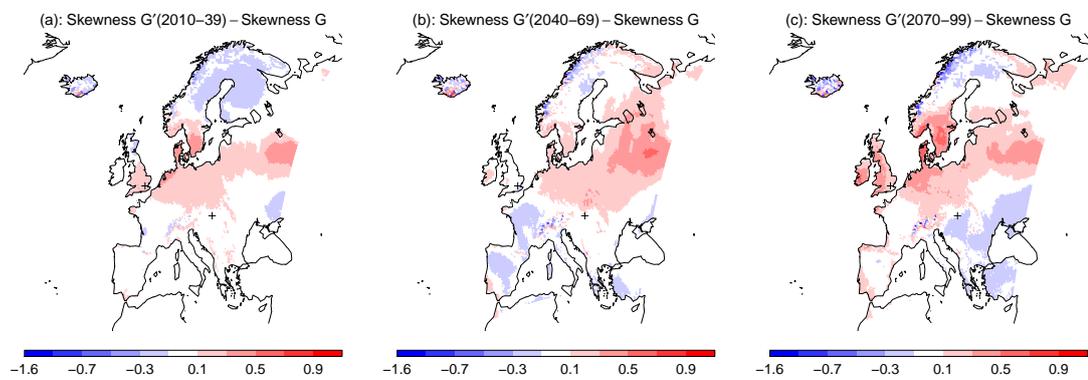


Figure 5.17: The differences between the sample moment skewness of G' for (a) 2010 to 2039, (b) 2040 to 2069 and (c) 2070 to 2099 and that of G .

5.6.4 Extreme temperatures calibrated by bias correction

The three bias correction methods, BC-L, BC-LS and BC-LSB, are used to calibrate HadRM3 temperatures. Estimates of location and scale parameters of BC-L and BC-LS

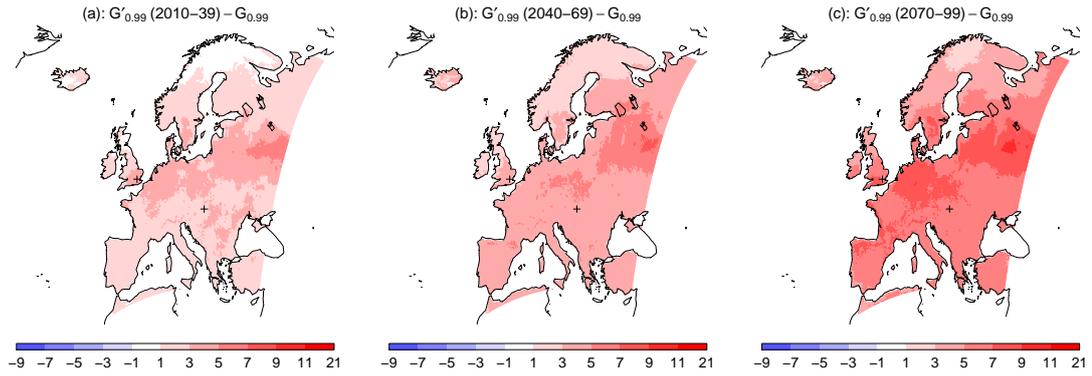


Figure 5.18: As in Fig. 5.17 but for differences in sample 0.99 quantile (in °C).

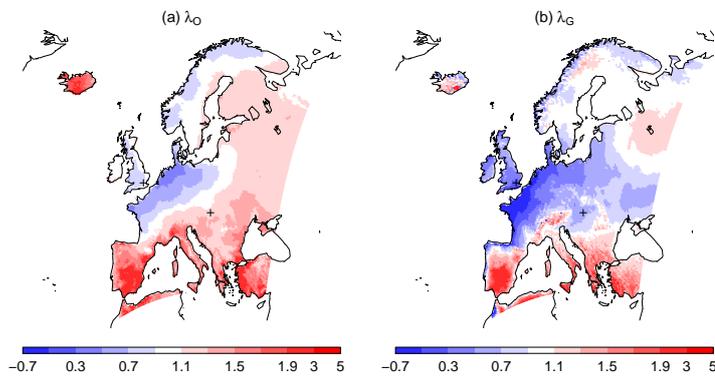


Figure 5.19: Maximum likelihood estimates of Box-Cox power transformation parameters λ_O and λ_G .

are the sample mean and standard deviation of O and G , which have already been shown in Section 5.6.1. The maximum likelihood estimates of power transformation parameters λ_O and λ_G for the BC-LSB method are shown in Figs 5.19(a) and (b) respectively. The spatial patterns shown on these maps associate closely with the maps showing the sample moment skewness of O and G [Figs 5.11(g) and (h)]. Areas with estimates of $\lambda < 1$ correspond to positive skewness, while areas with estimates of $\lambda > 1$ correspond to negative skewness. This is an expected result because as discussed in Section 5.2.2, to make a distribution to become more symmetric, values of $\lambda < 1$ are required for a positively skewed distribution, while values of $\lambda > 1$ are required for a negatively skewed distribution. These transformations allow the biases in the shape of G to be corrected.

As in Sections 5.4 and 5.5, the results to be presented here focus on the 10-summer

temperature return levels. Maps of extreme return levels for other return periods, say 5 years and 20 years, have similar spatial patterns. The biases in the 10-summer level of present-day HadRM3 temperatures calibrated by different bias correction methods (\hat{O}'_b) are shown in Fig. 5.20. Compared to the bias in the 10-summer level of uncalibrated temperatures in Fig. 5.12(e), BC-L can only reduce the warm bias slightly, with more than 5 °C of warm bias remaining over parts of central Europe, France and southeastern UK [Fig. 5.20(a)]. The use of BC-LS gives some improvement as the bias in the 10-summer level over central Europe is reduced to less than 1 °C [Fig. 5.20(b)]. However, a warm bias of about 4 °C still exists over southeastern UK and the northern coast of France and Spain. These are the regions where the distribution of G is more positively skewed compared to O [Fig. 5.11(i)]. Meanwhile, BC-LS appears to have over-corrected the bias over parts of Italy and southern Europe, where \hat{O}'_b has a negative bias of around 2 °C. In terms of the bias of \hat{O}'_b for the present-day, BC-LSB gives the best results. After this method is applied, the bias near the northern coast of France and Spain is reduced to less than 1 °C, and the warm bias in southeastern UK is reduced to about 2 °C. These results suggest that BC-LSB is able to correct the difference in shapes between the distributions of O and G . Since large biases remain after calibrating present-day HadRM3 temperatures using the BC-L method, only the maps for future temperatures calibrated by the other two methods (BC-LS and BC-LSB) are presented below.

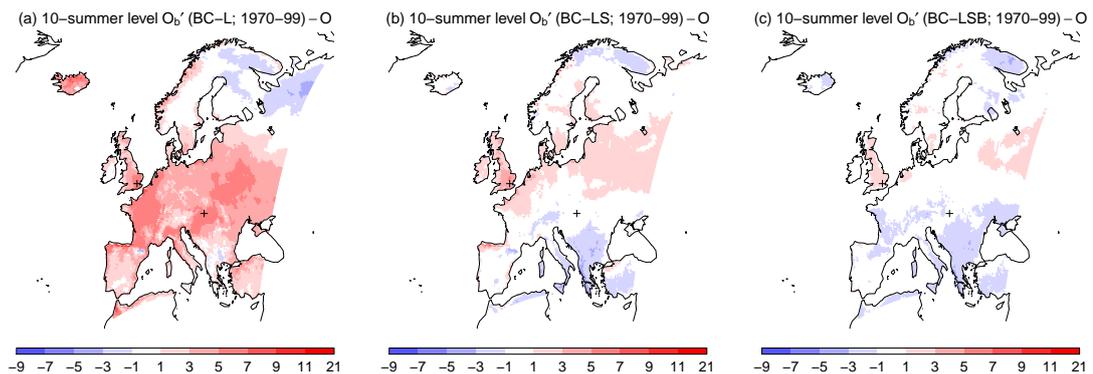


Figure 5.20: The difference (in °C) between the estimated 10-summer level of O'_b calibrated by (a) BC-L, (b) BC-LS and (c) BC-LSB for the present-day and the estimated 10-summer level of O .

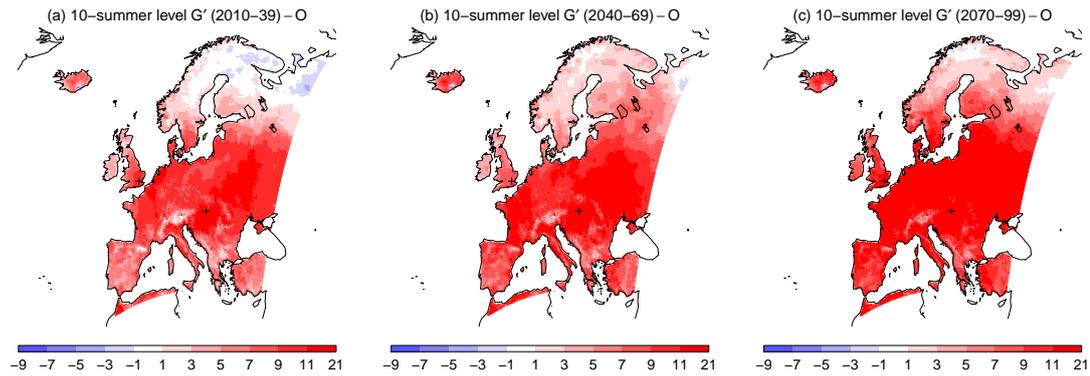


Figure 5.21: The difference between the estimated 10-summer level of uncalibrated temperatures G' for (a) 2010 to 2039; (b) 2040 to 2069 and (c) 2070 to 2099 and the estimated 10-summer level of O .

Figure 5.21 shows the future changes in 10-summer temperature return levels, if no calibration is performed. The largest increase in the 10-summer level occurs in France and central Europe. For example, in the period 2040 to 2069, the 10-summer level in these regions is about 10 °C higher than the observed level at the present-day [Fig. 5.21(b)]. A very different spatial pattern can be seen for temperatures calibrated by BC-LS shown in Figs 5.22(a) to (c). The largest increase in 10-summer level (which is about 6 °C in the period 2040 to 2069 compared to the present-day) occurs over northern Europe instead, including southern Scandinavia, northern Germany, the UK and western Russia. These are also the regions where the distribution of HadRM3 temperatures is projected to become more positively skewed (Fig. 5.17). In contrast, the projected rise in 10-summer return level over southern Europe is much reduced after the calibration.

Comparing Figs 5.22(a) to (c) and Figs 5.22(d) to (f), the main difference between the results for BC-LSB and BC-LS is that BC-LSB further lowers the estimated increase in 10-summer level over regions with a large difference in skewness between O and G , such as the UK and northern France. In the UK for example, the 10-summer level for temperatures in the period 2070 to 2099 calibrated by BC-LSB is about 4 to 6 °C higher than the present-day observed level [Fig. 5.22(f)]. The corresponding increase for temperatures calibrated by BC-LS is about 10 °C [Fig. 5.22(c)]. Nevertheless, after applying BC-LSB, the largest increase in the 10-summer level still occurs over northern Europe.

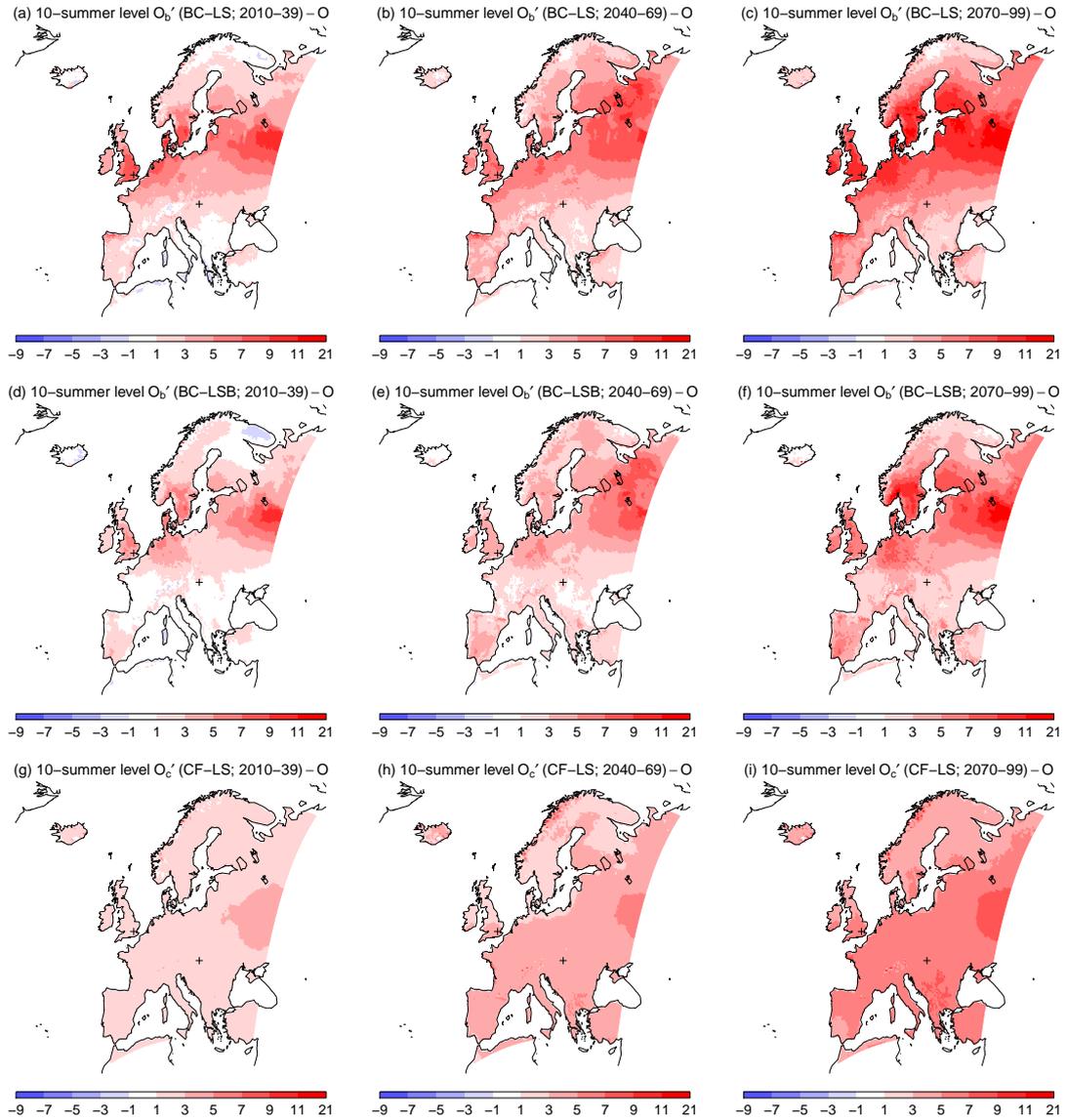


Figure 5.22: The difference (in °C) between the estimated 10-summer level of O' calibrated by BC-LS (a to c), BC-LSB (d to f) and CF-LS (g to i) for three periods (2010 to 2039; 2040 to 2069 and 2070 to 2099) and the estimated 10-summer level of O .

5.6.5 Extreme temperatures calibrated by change factor

It has been shown in Section 5.6.3 that the variance of G' is slightly greater than that of G towards the end of the century, therefore the assumptions of CF-LS method are considered to be more plausible than that of CF-L. The changes in 10-summer levels of temperatures calibrated by CF-LS relative to the observed present-day level are shown in Figs 5.22(g)

to (i), while the results for CF-L are omitted. Similar to the changes in model projected mean temperatures [Figs 5.15(d) to (f)], the increase in 10-summer level is rather spatially uniform, which is about 6 °C over continental Europe in the period 2070 to 2099. The increases over Scandinavia and Scotland are slightly less than that in other places.

5.6.6 Uncertainties of calibrated extreme temperatures

So far we have validated the assumptions for different temperature calibration methods by studying maps of sample statistics and calibrating the present-day model simulated temperatures G (for the bias correction approach only). Based on this work, BC-LSB and CF-LS are considered to be the most appropriate methods for calibrating HadRM3 temperatures over Europe². However, these two methods give substantially different results on future changes in extreme temperatures, particularly the spatial patterns of projected warming of high temperature return levels. For example, Fig. 5.23(a) shows the difference between the 10-summer level calibrated by BC-LSB and that calibrated by CF-LS for the period 2070 to 2099. The 10-summer level of \hat{O}'_b is generally about 4 to 6 °C lower than that of \hat{O}'_c over southern Europe, but is about 2 to 4 °C higher over Scandinavia and parts of UK and Germany. Such differences are possibly caused by a number of factors. First, the sample mean of \hat{O}'_b is about 2 °C lower than that of \hat{O}'_c over southern Europe [Fig. 5.23(b)], while there is little difference in their standard deviation [Fig. 5.23(c)]. Second, the 10-summer level of present-day temperatures calibrated by BC-LSB has a negative bias of about 2 °C compared to observations [Fig. 5.20(c)]. This might be carried over to future calibrated temperatures. These two factors could explain the difference between the 10-summer levels of \hat{O}'_b and \hat{O}'_c in southern Europe. As for the difference in the north, it is likely to be related to the change in skewness of modelled temperatures in the future over the region, as discussed in Section 5.6.3. Since the calibration by CF-LS does not incorporate changes in shapes between G and G' , the extremes of temperatures calibrated by this method are lower than that using BC-LSB.

²There might be different conclusions for individual locations, depending on the properties of observed and modelled temperature distributions. For example, in Section 5.5, BC-LS has been considered adequate for calibrating temperatures in Budapest.

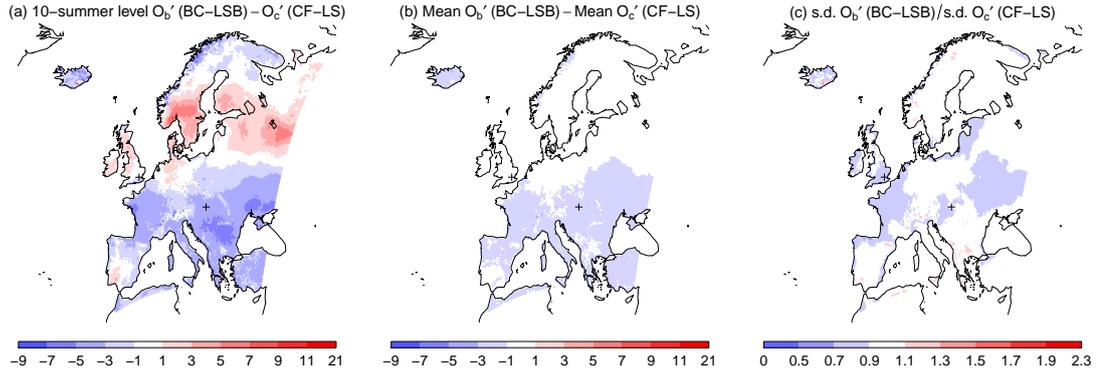


Figure 5.23: (a) Difference in the 10-summer level (in °C), (b) difference in the sample mean (in °C) and (c) ratio of the standard deviation between \hat{O}'_b calibrated by BC-LSB and \hat{O}'_c calibrated by CF-LS for the period 2070 to 2099.

As shown in Section 5.2.4, the mean of O'_b and the mean of O'_c could be different even if the distributional assumptions for both calibration approaches are satisfied. It is therefore not unreasonable to expect extreme temperatures calibrated by the two approaches to be different. This calibration uncertainty is an additional source of uncertainty of climate model projections.

Another source of uncertainty which has been considered in this chapter is the uncertainty associated with physical parameters of the RCM (see Section 2.3.3). For London and Budapest, this uncertainty has been assessed by considering the calibrated projections from 11 PPE members of HadRM3. In order to compare the relative contribution of uncertainty from the use of different PPE members with the uncertainty from the choice of different calibration approaches, the estimated 10-summer levels of \hat{O}'_c for the 11 PPE members are plotted against the corresponding estimates for \hat{O}'_b for the period 2070 to 2099 in Fig. 5.24. The vertical distance of each data point from the line of equal values (dashed line) represents the uncertainty of the choice of calibration method, while the spread of data points on each axis represents the uncertainty of climate model parameters. For London [Fig. 5.24(a)], the mean difference between the 10-summer temperature level estimates for each PPE member using the two calibration approaches is about 3 °C, while the ranges of estimates for BC-LSB and CF-LS are about 7 °C and 3 °C respectively. For Budapest [Fig. 5.24(b)], the mean difference between the estimates using the two calibra-

tion approaches is around 3 °C, and the ranges of estimates for BC-LS and CF-LS are about 4 °C and 5 °C respectively. These comparison suggest that calibration uncertainty and climate model parameter uncertainty are comparable. Their effects on mortality projections will be examined in Chapter 6. Although temperature projections from the PPE for the rest of Europe are outside the scope of this thesis, one should be aware of this uncertainty of climate model parameters when interpreting the extreme calibrated temperatures from different methods.

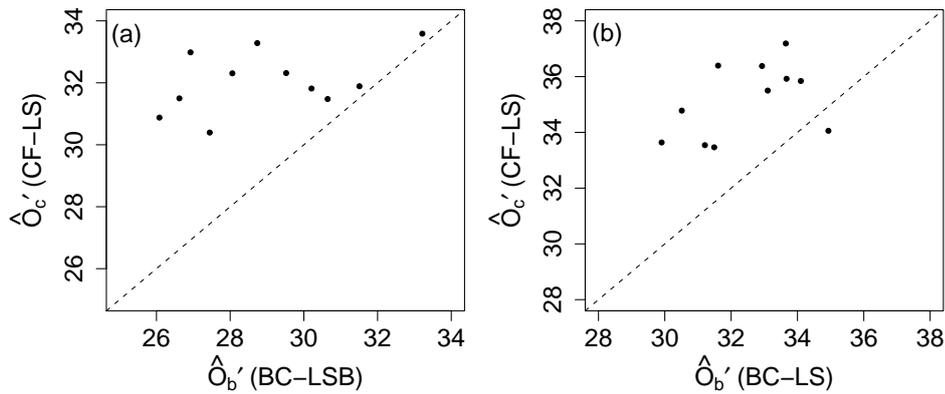


Figure 5.24: Scatter plot of estimated 10-summer levels of 11 PPE members for (a) \hat{O}'_c (in °C) calibrated by CF-LS against \hat{O}'_b (in °C) calibrated by BC-LSB in London and (b) \hat{O}'_c (in °C) calibrated by CF-LS against \hat{O}'_b (in °C) calibrated by BC-LS in Budapest for the period 2070 to 2099.

5.6.7 Limitations of calibration methods

We finally consider the limitations of the calibration methods adopted in this chapter. Possible future work to address these limitations will be discussed in Chapter 7. First, all the methods discussed in this chapter are designed to calibrate the entire temperature distribution. Changes in extreme temperatures are estimated after the calibration. While these methods might be able to calibrate the discrepancies between the tails of modelled and observed temperatures in some cases, there is still room for improvement. For example, even though BC-LSB considerably reduces the bias of G at the extreme compared to other bias

correction methods, the mapping of modelled temperature values onto observed temperature values by BC-LSB for London as shown by the red line in Fig. 5.3(a) still indicates discrepancies between the calibration and the actual relationship for the sample quantiles of O and G . It will be worth exploring alternative calibration techniques which specialize on the extremes of the distributions.

Second, an important assumption in the bias correction approach is that through the use of a constant transfer function, the bias of the model does not change with time. It is not possible to validate this assumption.

Third, a ‘time-slice’ approach is adopted in this calibration exercise. The set of parameters for the calibration, such as sample mean and standard deviation, are calculated from temperature samples in 30-year periods. We have assumed that the observed and modelled daily temperatures in each time-slice are independent and identically distributed. With the presence of annual cycles and possible long-term trends in the location, scale and shape of temperatures, the validity of this assumption can be questionable.

5.7 Summary

This chapter has considered two generic approaches to calibrate HadRM3 temperature projections. For each approach, a number of methods have been assessed. These are based on assumed relationships of properties between the distributions of observed and modelled temperatures. In particular, a novel method involving the Box-Cox transformation has been used to correct the bias in the shape of the distribution of modelled temperatures. This method has considerably reduced the bias in the present-day extreme temperatures in London simulated by HadRM3. For the future, the two calibration approaches give substantially different estimates of extreme temperatures at certain locations, but they are significantly lower than the estimates projected by HadRM3 without calibration. For London and Budapest, temperature projections from the perturbed physics ensemble of HadRM3 have been calibrated. The spread of extreme calibrated temperatures among the ensemble members is comparable to the difference of extreme temperatures calibrated by the two approaches.

Chapter 6

Extreme heat-related mortality projections

6.1 Aim

This chapter presents the projection of extreme heat-related mortalities in London and Budapest using the mortality models developed in Chapter 4, driven by the calibrated HadRM3 projected temperatures presented in Chapter 5. The uncertainties of mortality projections are assessed by considering the sensitivity of projections to the choice of mortality model, the different temperature projections from HadRM3 ensemble members and the choice of temperature calibration methods.

6.2 Projection methodologies

This section describes how the data and results from previous chapters are used to project future heat-related mortalities in London and Budapest. We first revisit the systematic components of the two statistical models developed separately for the two locations in Chapter 4:

$$\text{Model TO: } \log \mu_i = \log P_i + f_1(r_i) + f_5(T_i^o);$$

$$\text{Model PTO: } \log \mu_i = \log P_i + f_1(r_i) + \beta_3 T_i^o + \beta_4 (T_i^o)^2,$$

where μ_i is the mean elderly mortality rate, P_i is the daily elderly population estimate, r_i represents time, and T_i^o are daily mean temperatures (refer to Section 4.4 for details). In both models, the first offset term on the right hand side represents changes in elderly population, while the second term captures the variation in daily elderly mortality not related to temperatures, including its annual cycles and long-term trend. The dependence of daily elderly number of deaths on gridded daily mean air temperature is represented by the third term on the right hand side of model TO as a non-parametric smooth function, and by the third and fourth terms on the right hand side of model PTO as a second order polynomial. With estimated $f_1, f_5, \beta_3, \beta_4$, future population and calibrated temperature projections from HadRM3, future summer daily elderly mortalities for these two cities can be estimated. As in the calibration of future RCM temperature projections in Chapter 5, daily elderly number of deaths are projected for three 30-year periods: 2010 to 2039, 2040 to 2069 and 2070 to 2099. The extreme elderly mortalities for each period are then assessed.

6.2.1 Assumptions on population and mortality annual cycles

As noted in Section 4.4, the mortality models in this thesis are specified in a different way compared to the models in previous mortality projections reviewed in Section 2.5. The response of the models in previous studies is typically the daily excess (elderly) mortality. This is calculated by subtracting the baseline mortality from the observed mortality counts, such that the variations in population and non-temperature related mortality are removed prior to model fitting (refer to Section 4.2.4 for details). On the other hand, the response of models TO and PTO is the total elderly mortality counts. Two explanatory variables, elderly population estimates (P_i) and time (r_i), are used in the models to represent baseline variations. This different specification gives projections in the unit of number of deaths per day which are easier to be interpreted.

In the mortality projections to be presented in Section 6.3, elderly population is assumed to be unchanged from the estimates on 15th July 2003 for London and on 15th July 2001 for Budapest, i.e. mid-summer in the last year of the observed mortality time series for each

location. These projections then reflect the impacts of climate change on changes in future mortalities. The additional effects of demographic changes on mortalities are assessed as part of the sensitivity analyses to be described in Section 6.2.2. In addition, the mortality annual cycles in the future are assumed to be identical to that in 2003 for London and in 2001 for Budapest. For example, let $\hat{T}_{j,k}^{\prime}$ be the calibrated temperature for London on day j in the future which is also the k^{th} day in the summer of that year. The corresponding number of elderly deaths by model TO $\hat{\mu}_{j,k}$ is calculated using

$$\hat{\mu}_{j,k} = \exp[\log P_{3848} + \hat{f}_1(r_{3785+k}) + \hat{f}_5(\hat{T}_{j,k}^{\prime})],$$

where P_{3848} is the interpolated elderly population size in London on 15th July 2003, r_{3785+k} is the time in years from the start date of the observed mortality time series (1st January 1993)¹, and \hat{f}_1 and \hat{f}_5 are the estimated smooth functions of model TO for London.

6.2.2 Sensitivity analyses

Uncertainties in mortality projections arise from uncertainties in climate projections, uncertainties in modelling present-day daily mortalities and uncertainties in the future changes of population characteristics. These are shown schematically in Fig. 6.1. Below briefly discusses each of these sources of uncertainties and describes how the sensitivity of mortality projections to these uncertainties is assessed.

¹15th July 2003 is on the 3848th day of the London population time series, and the start date of 2003 summer (15th May) is on the 3786th day of this series. Refer also to Section 4.4.

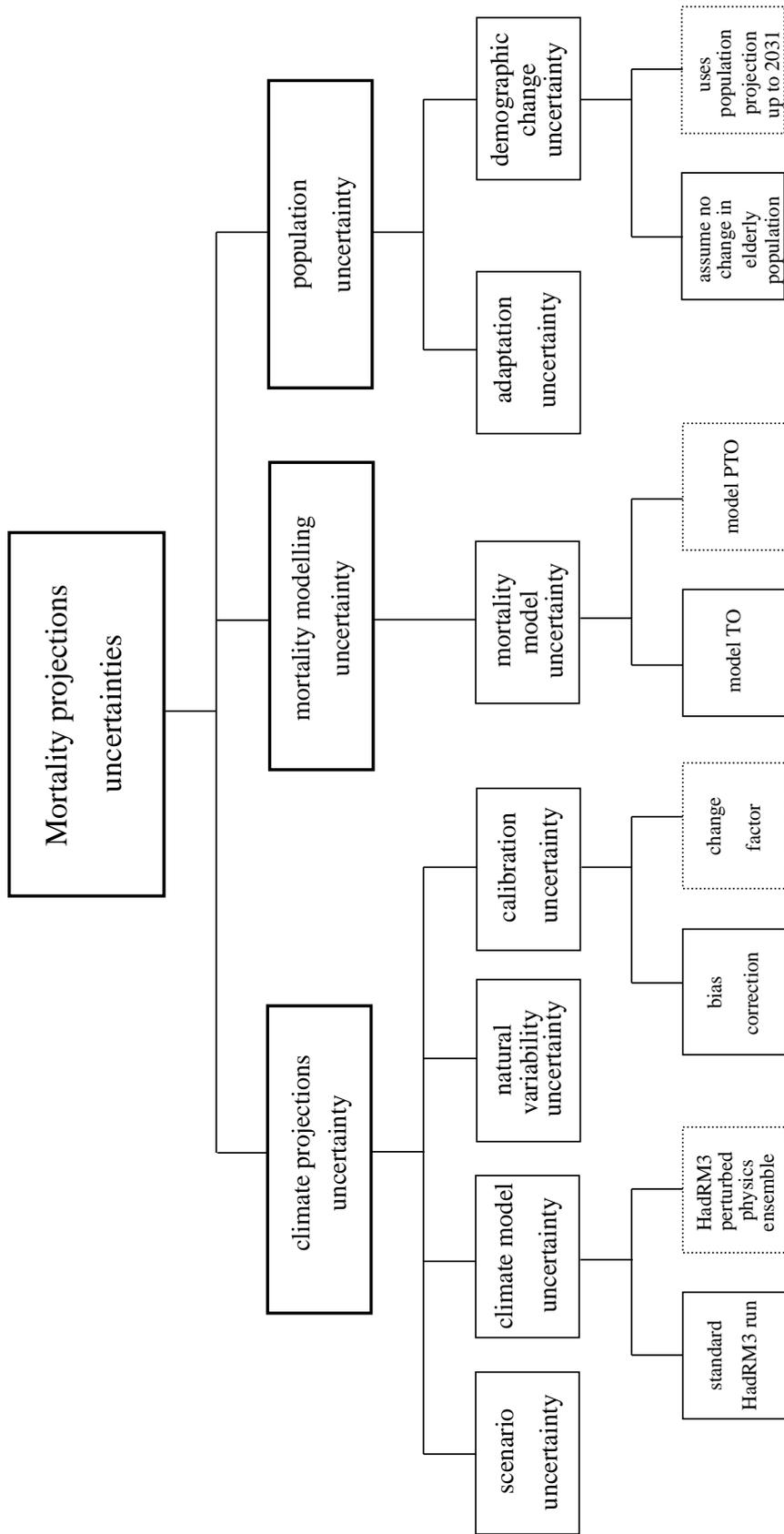


Figure 6.1: Schematic diagram showing sources of uncertainties in mortality projections. The bottom row shows how the sensitivity of four of these sources uncertainties to extreme mortality projections are investigated in this thesis. The results of projection using the data, mortality model and calibration method in the solid boxes of this row are presented in Section 6.3, while the results for the dotted boxes are presented in Section 6.4.

(a) Climate projections uncertainty

As discussed in Section 2.3.3, uncertainties in future emissions of greenhouse gases (scenario uncertainty), representation of the climate system using climate models (climate model uncertainty) and natural climate variability contribute to the overall uncertainties in climate projections. Since the HadRM3 simulations used in this thesis are driven by a *single* emission scenario (SRES A1B; refer to Section 3.4), the effects of scenario uncertainty on mortality projections are not considered here. For climate model uncertainty, the 11-member PPE of HadRM3 is used to evaluate the uncertainty related to the choice of RCM parameters (parameter uncertainty). This is done by first projecting future mortalities based on calibrated temperature projections from each PPE member. The range of extreme mortalities projected for the 11 members then represents the impacts of parameter uncertainty on extreme mortality projections. As for the uncertainty arising from natural climate variability, this is not considered in this sensitivity analysis as the PPE of HadRM3 is not designed to systematically sample this uncertainty.

In addition, the results from Chapter 5 suggest that the choice of RCM temperature calibration method contributes a further source of uncertainty to the projections of extreme temperatures. In order to study the impact of this uncertainty on mortality projections, extreme mortalities projected using bias-corrected HadRM3 temperatures (BC-LSB for London; BC-LS for Budapest) are compared with projections using temperatures calibrated by the change factor approach (CF-LS for both locations).

(b) Mortality modelling uncertainty

The second category of uncertainty is related to the predictions and specification of the mortality model. The uncertainty in model predictions associated with model parameters can be estimated by constructing confidence intervals of the predicted mean response, as discussed in Section 4.3.1. As for model specification, there are uncertainties in the choices of covariates and functional dependence of mortality on the covariates (see Section 4.2). Specifically for the mortality models in this thesis, there are uncertainties in the choice of the degree of smoothing in the non-parametric function representing the mortality annual cycles, i.e. f_1 in models TO and PTO (see Section 4.4), and how the relationship between daily mortality and temperature is modelled. It has been shown in Section 4.5.3

that modelling the mortality-temperature relationship with a non-parametric smooth function (model TO) and modelling such a relationship with a second order polynomial (model PTO) give different predictions for the upper tail of the mortality distribution for London. The sensitivity of extreme mortality projections to this particular specification of mortality model is assessed.

(c) Population uncertainty

A further category of uncertainty is related to future changes in population (demographic change uncertainty) and its response to extreme heat (adaptation uncertainty). The future number of heat-related deaths will obviously depend on the size of the susceptible group of population, in this case the number of elderly people. With an ageing population generally over Europe (Chapter 1 in Eurostat, 2008), this is a particularly important factor to consider. Gosling et al. (2009a) suggests the possible use of SRES global population growth scenario (Nakićenović and Swart, 2000) to evaluate the impacts of demographic changes on future heat-related mortalities. This is not considered here because these global projections cannot reflect any regional population changes, for example regional differences in fertility and mortality rates and migration. Instead, this thesis uses the annual elderly population projection data for London and Budapest up to the year 2031 described in Section 3.2.2. Despite covering a relatively short period of time, these projections have incorporated the above-mentioned regional changes and are therefore considered to be more precise. For this sensitivity analysis, the projected daily elderly population from 2010 to 2030 are obtained by linearly interpolating the annual elderly population projections. Mortalities are projected for this 21-year period with the daily elderly population projections as P_i in the mortality model. These are then compared with the projections which keep P_i at the mid-2003 level (as described in Section 6.2.1). However, it should be noted that, being projections based on current demographic trends, the population projections used here are themselves subject to uncertainties. This is especially the case for Budapest where the population projection is derived from projections for a larger geographical region (see Appendix A).

Adaptation uncertainty refers to uncertainties in how the population and the society will acclimatize to future climates in the long-term, as discussed in Section 2.4.3. Carson

et al. (2006) estimated the relationships between weekly number of deaths and temperatures in London for four 11-year periods from 1900 to 1996. The mortality-temperature gradient for heat-related deaths was found to decrease over the century. This gives some empirical evidence of changing mortality-temperature relationship over time in the past. If the population and the society are able to adapt to warmer temperatures in the future, the modelled mortality-temperature relationship based on present observations may overestimate future heat-related mortalities. However, adaptation is difficult to be modelled because this depends not only on human physiology, but also on other factors such as the changes in the provision of medical care services and the wealth of the society (Ebi, 2008; McGregor et al., 2007).

As discussed in Section 2.5, two different approaches have been adopted by previous mortality projections to assess the effects of adaptation. One of them is the ‘analogue city’ approach adopted by Kalkstein and Greene (1997). This approach projects the future heat-related mortality at one city using the present-day mortality-temperature relationship of an analogue city. Since the future climate of the analogue city is similar to the present climate of the city of interest, it is assumed that the mortality-temperature relationship in the city of interest will change to the present-day observed relationship for the analogue city. Another approach, adopted by Dessai (2002, 2003) and Gosling et al. (2007, 2009b), considers simple adjustments to the modelled present-day mortality-temperature function. The mortality-temperature function is assumed to translate with time to represent adaptation. Consider our model TO as an example, if the daily mean temperature on a future day is $T_i^{o'}$, the estimated number of deaths will be

$$\hat{\mu}_i = \exp \left[\hat{f}_5 \left(\hat{T}_i^{o'} - c(r_i) \right) + \text{other non-temperature related terms} \right],$$

where c depends on time r_i . This is shown schematically in Fig. 6.2. If a day with extremely high temperature at the present repeats in the future, fewer number of deaths will be projected compared to the expected number at the present. The choice of c was arbitrary. Dessai (2003) assumed a 1 °C translation for every 30 years, while Gosling et al. (2009b) examined the sensitivity of mortality projections in the period 2070 to 2099 to

translations of 2 °C and 4 °C relative to the late 20th century². As noted by Gosling et al. (2009b), there is no evidence suggesting how the mortality-temperature function will change in the long-term. It is also uncertain whether a simple translation of this function is a reasonable assumption. Therefore, rather than making crude assumptions regarding adaptation, the effects of this uncertainty are not assessed in this thesis. All the mortality projections to be presented in this chapter are adaptation independent, i.e. the mortality-temperature function is assumed to be stationary.

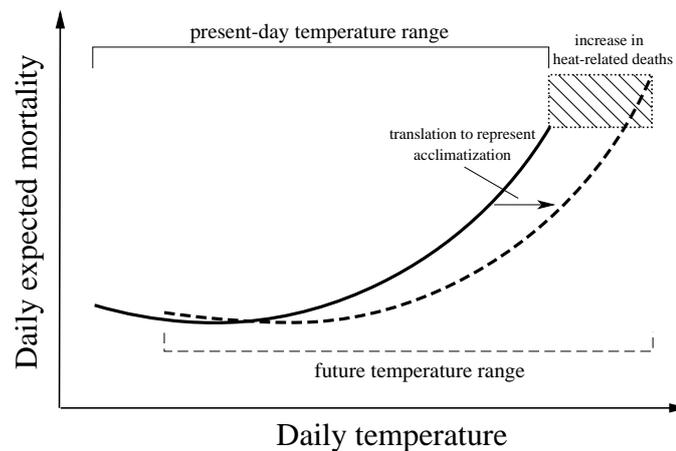


Figure 6.2: Schematic representation of possible future changes in the summer mortality-temperature relationship under a warming situation. The relationships for the present-day and a certain time in the future are shown by thick solid and dashed curves respectively. The hatched box indicates the increase in number of heat-related deaths. Adapted from Fig. 2 of McMichael et al. (2006).

The four sensitivity analyses to be carried out in this thesis are shown in the bottom row of Fig. 6.1. In order to make a more systematic comparison, Section 6.3 first presents the results for the mortality projections using

- mortality model TO;
- temperatures projected by the standard run of HadRM3 which are calibrated using the bias correction approach (BC-LSB for London; BC-LS for Budapest) and
- constant elderly population from mid-summer 2003.

²Gosling et al. (2007) assessed mortality projections for 6 different cities. The mortality and temperature data which are used to fit the mortality models cover different time periods in the late 20th century.

Then in Section 6.4, these results are compared with mortality projections using alternative choices of mortality model, temperature calibration approach, HadRM3 PPE projections and population projection data. Two important points regarding these sensitivity analyses need to be noted. First, these analyses do not constitute a complete assessment of uncertainties in mortalities because some sources of uncertainties are explicitly excluded, such as scenario uncertainty and adaptation uncertainty. Even for climate model uncertainty which is considered in these analyses, the HadRM3 PPE can only represent the uncertainty in the RCM parameters, but not uncertainties in the structural choices of the model (see Section 2.3.4). Second, different sources of uncertainties are not necessarily independent. For example, future emission of greenhouse gases is related to future changes in population (O'Neill, 2004).

6.2.3 Estimation of return levels of extreme mortality counts

As mentioned in Section 4.3.1, the use of GP distribution is not applicable on modelling extreme mortality counts. The extreme mortalities for the present-day (1993 to 2003 for London; 1992 to 2001 for Budapest) and each of the three 30-year future periods are therefore estimated using the sample quantiles of observed and projected mortality counts in each period. Unlike the assessment of extreme calibrated temperatures in Chapter 5 which focuses on the 10-summer return level, the discussion here focuses on 2-summer mortality return level, i.e. the highest number of daily elderly deaths expected to occur every other year on average. This is because with fewer data values at the tail of the distribution of projected mortality, the estimation of rarer extreme number of deaths using sample quantiles may be less precise.

6.3 Projections of future extreme summer mortalities

6.3.1 Projections for London

Table 6.1 shows the estimated mean and 2-summer return levels of projected number of daily elderly deaths in London, using mortality model TO and HadRM3 projected tem-

peratures calibrated by BC-LSB. The estimates for present-day (1993 to 2003) observed mortality are also shown for comparison. Although the mean of calibrated daily temperatures in the three future periods are all higher than the mean of present-day observed daily temperatures, the mean of projected number of deaths in the future are lower. This can be explained by the decreasing trends of both the elderly population and the number of deaths throughout the present-day period [refer to Figs 3.1(c) and 3.3(a)]. Since future mortalities are projected with the assumption of constant population from mid-summer 2003 (near the lowest level in the time series), the mean of projected number of daily deaths in the future can be lower than that at the present. With a projected increase of mean temperatures, the mean number of daily deaths is projected to increase slightly.

Time period	Mean		2-summer return level	
	Temperature (°C)	No. of deaths	Temperature (°C)	No. of deaths
1993-2003	16.6	121	25.0	170
2010-2039	17.4	114	27.3	193 [178,208]
2040-2069	18.4	116	26.3	175 [165,185]
2070-2099	19.3	120	28.5	218 [196,240]

Table 6.1: The projected mean and 2-summer return level of temperature ($\hat{T}^{o'}$) and number of elderly deaths ($\hat{\mu}$) in London, using mortality model TO, temperature projections from the standard run of HadRM3 calibrated by BC-LSB and assuming future elderly population to be constant at the mid-summer 2003 level. The 95% confidence intervals of number of deaths predicted by model TO are given in square brackets (see Section 4.2 for how these are estimated). The observed mean and 2-summer levels of temperature (T^o) and number of deaths (Y) in the period 1993 to 2003 are also shown.

While the increase in the 2-summer temperature level is similar in magnitude to the increase of mean temperatures, the 2-summer mortality return level is projected to increase more rapidly than the increase in the mean daily number of deaths. The estimated 2-summer level in the period 2070 to 2099 is about 28% higher than that at the present. The more rapid increase in future extreme mortalities is related to the non-linear relationship between daily number of deaths and temperatures for this city [refer to Fig. 4.4(c)].

6.3.2 Projections for Budapest

Due to the decreasing trends of population size and observed number of deaths [refer to Figs 3.2(c) and 3.3(b)], the projected mean and 2-summer return levels of mortality in Budapest are also lower than that in the present-day period (1992 to 2001), as shown in Table 6.2. The projected increase in mean and 2-summer return levels of temperatures calibrated by BC-LS are similar in magnitudes - about 3 °C from the present to the period 2070 to 2099. The mean number of daily elderly deaths is projected to increase by about 8% between the periods 2010 to 2039 and 2070 to 2099, while the 2-summer mortality level is projected to increase by about 12%.

Time period	Mean		2-summer return level	
	Temperature (°C)	No. of deaths	Temperature (°C)	No. of deaths
1992-2001	20.5	51	28.9	77
2010-2039	21.2	48	30.4	65 [60,70]
2040-2069	22.3	50	31.1	68 [62,74]
2070-2099	23.4	52	32.1	73 [66,80]

Table 6.2: Same as Table 6.1 but for Budapest. The mortality projections are based on mortality model TO and temperature projections from the standard run of HadRM3 calibrated by BC-LS, assuming future elderly population to be constant at the mid-summer 2001 level. The observed mean and 2-summer level of T^o and Y are for the period 1992 to 2001.

6.4 Sensitivity analyses

6.4.1 Climate projections uncertainty

Here we first consider the climate model uncertainty and then the calibration uncertainty. As discussed in Sections 5.4.3 and 5.5.3, there is considerable spread among the future changes in extreme temperatures projected by different members of the PPE. This leads to

large differences in projected extreme mortalities. Consider, for example, mortality projections for London in the period 2070 to 2099 using model TO and HadRM3 temperatures calibrated by BC-LSB. The range of 2-summer temperature return level estimates among the PPE members is 5.9 °C [Fig. 6.3(a)]. The corresponding 2-summer mortality level estimates range from 162 to 289 (ensemble mean: 205) [Fig. 6.4(a)]. To put this large spread into context, the 2-summer mortality level for 2 of the 11 PPE members are higher than the maximum daily number of deaths observed during the 2003 heatwave (233), which is shown by the blue dashed line in Fig. 6.4(a). These 2 members project that this high daily number of deaths will occur at least every other year on average in the period 2070 to 2099. Meanwhile, there are 2 PPE members which project the 2-summer mortality level to be lower than the present-day 2-summer level, which is shown by the blue solid line in Fig. 6.4(a).

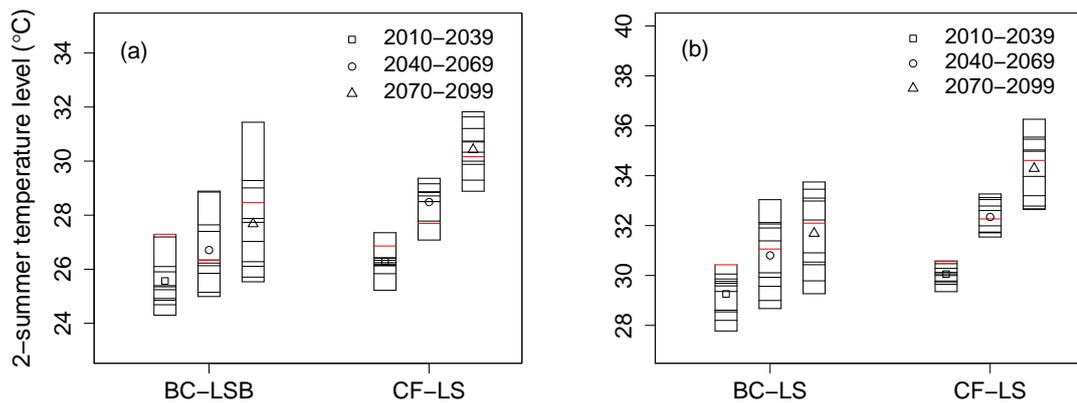


Figure 6.3: Estimates of 2-summer return levels of calibrated HadRM3 temperature projections for (a) London and (b) Budapest for each member of PPE, using different calibration methods as labelled. The symbols in each box represent the mean of the return levels of the PPE members for each future 30-year period as labelled. The black horizontal lines within each box show the return level estimates of each PPE member, with the estimate for the standard run shown in red.

In comparison, the range of 2-summer mortality levels among the PPE members is smaller for Budapest using model TO and temperatures calibrated by BC-LS [Fig. 6.4(b)]. This can be explained partly by the narrower range in the 2-summer temperature levels

[Fig. 6.3(b)]. In addition, the expected number of deaths predicted by model TO increases less rapidly with extreme temperatures in Budapest compared to London [compare Figs 4.4(c) and (d)]. The 2-summer mortality level for Budapest is therefore less sensitive to different extreme temperature projections by the PPE members.

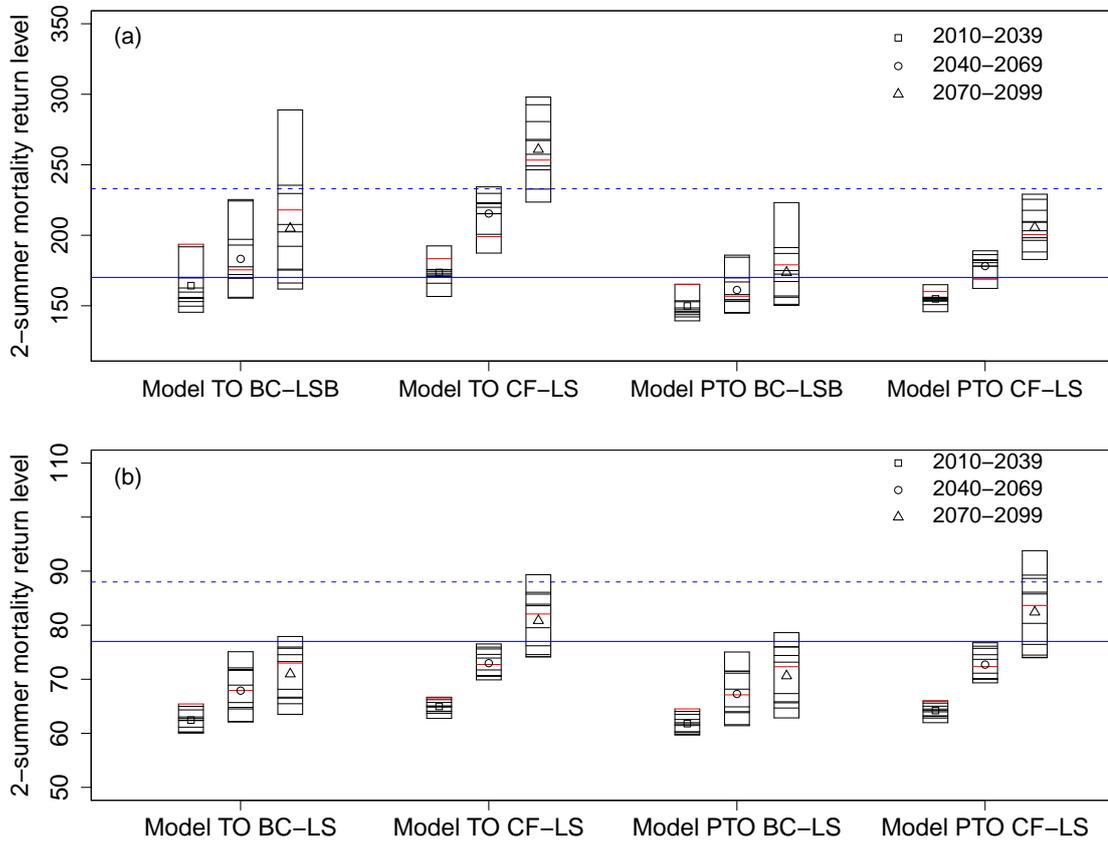


Figure 6.4: Estimates of 2-summer mortality return level in (a) London and (b) Budapest by HadRM3 PPE members using different combinations of mortality model and temperature calibration method as labelled. The blue solid line represents the 2-summer return level estimate for observed mortality in 1993 to 2003 for London and in 1992 to 2001 for Budapest. The blue dashed line indicates the maximum daily number of deaths in the period. Refer to the caption of Fig. 6.3 for explanation of other symbols.

Comparing the length of the first three boxes on each panel of Fig. 6.3, it can be seen that for both locations, the ensemble spread of 2-summer temperature levels increases with time. This suggests increasing climate model uncertainties in extreme temperature

projections. As a result, the associated uncertainty in extreme mortality also increases with time. For example, the range of 2-summer mortality level estimates for London using model TO and BC-LSB increases from 44 in the period 2010 to 2039 to 127 in the period 2070 to 2099 [Fig. 6.4(a)].

(a) London				
Time period	Model TO	Model TO	Model PTO	Model PTO
	BC-LSB	CF-LS	BC-LSB	CF-LS
2010-2039	164	174	150	155
2040-2069	183	215	161	178
2070-2099	205	261	174	206
(b) Budapest				
Time period	Model TO	Model TO	Model PTO	Model PTO
	BC-LS	CF-LS	BC-LS	CF-LS
2010-2039	62	65	62	64
2040-2069	68	73	67	73
2070-2099	71	81	71	82

Table 6.3: Ensemble mean of 2-summer mortality return level estimates using different combinations of mortality models and temperature calibration methods for (a) London and (b) Budapest.

The calibration uncertainty is assessed by comparing the 2-summer mortality level projected by model TO using temperatures calibrated by bias correction (BC-LSB for London; BC-LS for Budapest) and that using temperatures calibrated by change factor (CF-LS). For both locations, the ensemble mean 2-summer mortality level estimates using temperatures calibrated by CF-LS are higher than that using temperatures calibrated by bias correction (Table 6.3). In the period 2070 to 2099 for example, the estimated ensemble mean 2-summer level of number of deaths in London using CF-LS is 27% greater than that using BC-LSB. Among the 11 PPE members, 10 of them project that the highest daily number of deaths in the 2003 heatwave will occur at least every other year on average [Fig. 6.4(a);

c.f. only 2 members for BC-LSB]. For the same period in Budapest, the ensemble mean 2-summer mortality level using CF-LS is 14% greater than that using BC-LS. These differences are due to the higher ensemble mean 2-summer level of temperatures calibrated by CF-LS (Fig. 6.3).

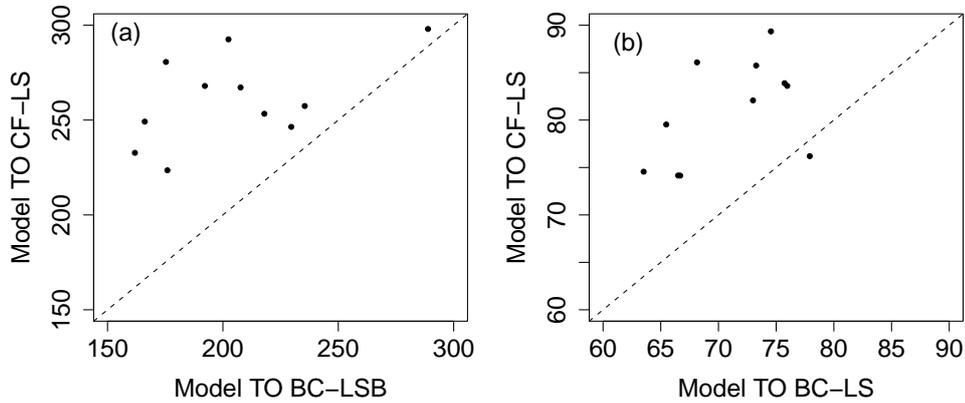


Figure 6.5: Scatter plot of estimated 2-summer mortality level projected by model TO for 11 PPE members in the period 2070 to 2099 using (a) $T^{o'}$ calibrated by CF-LS against that using $T^{o'}$ calibrated by BC-LSB for London and (b) $T^{o'}$ calibrated by CF-LS against that using $T^{o'}$ calibrated by BC-LS for Budapest.

The relative impacts of climate model uncertainty and calibration uncertainty on extreme mortality projections can be compared using Fig. 6.5. Similar to the scatter plots shown in Section 5.6.6, Fig. 6.5 plots the estimated 2-summer mortality level in the period 2070 to 2099 for the 11 PPE members using bias-corrected temperatures against that using temperatures calibrated by change factor. For London [Fig. 6.5(a)], the mean difference between the 2-summer mortality level estimates for each PPE member using the two calibration approaches (vertical distance of each data point from the dashed line of equal values) is around 55. Excluding the outlier on the far right, the range of 2-summer mortality level estimates among the 11 members for each calibration approach (range of values on each axis) is about 70. For Budapest [Fig. 6.5(b)], the mean difference between the estimates for each PPE member using the two calibration approaches and the range of estimates among the PPE members are both about 10. This comparison suggest that climate model uncertainty and calibration uncertainty have comparable effects on extreme

mortality projections in this time period. Plots for other time periods (not shown) also support this conclusion.

6.4.2 Mortality modelling uncertainty

The mortality model which specifies the mortality-temperature relationship as a second order polynomial (model PTO) gives lower estimates of extreme mortality in London than the model which specifies such a relationship as a non-parametric smooth function (model TO). Using temperatures calibrated by BC-LSB, the ensemble mean 2-summer mortality level for model PTO is 9 to 15% lower than that for model TO [Table 6.3(a)]. The difference between the ensemble mean 2-summer mortality level for the two mortality models using temperatures calibrated by CF-LS is larger (11 to 21%). This is because the difference between the predicted number of deaths of the two mortality models increases with temperature [refer to Figs 4.7(a) and (b) for a comparison between the predictions of two similar models, TS and PTS³], and ensemble mean 2-summer levels of temperatures calibrated by CF-LS are higher than that calibrated by BC-LSB (Fig. 6.3). Figure 6.6(a) compares the 2-summer mortality levels in London projected by the two mortality models in the period 2070 to 2099 for the 11 PPE members. The mean difference between the estimates for each PPE member using the two mortality models is about 30, which is less than the range of ensemble estimates projected by model TO. This suggests that climate model uncertainty has a greater effect on the 2-summer mortality return level projections than the mortality model uncertainty. However, it is possible that the mortality model uncertainty will become more important for even rarer extremes of daily mortality counts as the difference between the predictions by the two mortality models further increases with temperature.

In contrast, for Budapest, there is little difference between the ensemble mean 2-summer mortality level projected by the two mortality models [Table 6.3(b)]. This is also true for individual PPE members. An example is given in Fig. 6.6(b) which shows that in the pe-

³The only difference is that models TS and PTS use station observed temperatures T^s , while models TO and PTO use gridded observed temperatures T^o . Refer to Section 4.4 for details.

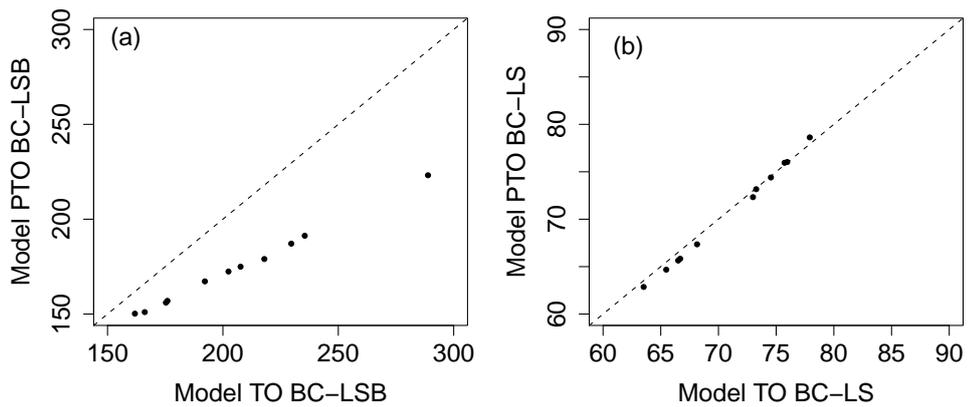


Figure 6.6: Scatter plot of estimated 2-summer mortality level for 11 PPE members in the period 2070 to 2099 projected by model TO against that projected by model PTO for (a) London (using temperatures calibrated by BC-LSB) and (b) Budapest (using temperatures calibrated by BC-LS).

riod 2070 to 2099, the 2-summer mortality level estimates for each PPE member using model TO agree well with the corresponding estimates using model PTO. This is because there are little differences between predictions by the two models at extreme temperatures [refer to Figs 4.7(c) and (d)].

6.4.3 Population uncertainty

The mortality projection results discussed so far have assumed the future elderly population to be constant. Table 6.4 shows how the mortality projections differ when elderly population projection data are used. The results shown are based on mortality projections using model TO and bias-corrected HadRM3 temperatures for the time period 2010 to 2030, because population projections for both locations are only available up to the year 2031. For London where the elderly population size is projected to increase steadily [refer to Fig. 3.3(a)], the ensemble mean of mean number of deaths is about 7% higher than the present-day mean number of deaths, while the ensemble mean 2-summer mortality level is about 15% higher than the present-day 2-summer level (refer to Table 6.1). In comparison, for projections for the same time period using the constant population as-

sumption, the ensemble mean of mean and 2-summer level are about 7% and 4% lower than the present-day respectively. These suggest that relative to warming in temperatures, the changes in elderly population size have a possibly greater effect on changes in future elderly mortalities in this period. However it is not possible to say whether this will remain true later in the century, given the uncertainty in the future population trend and more importantly, the increasing uncertainties in mortality projections related to uncertainties in climate projections.

Location	(a) assumes constant population		(b) uses projected population	
	Mean	2-summer return level	Mean	2-summer return level
London	113	164	130	195
Budapest	48	62	52	68

Table 6.4: Ensemble mean of mean and 2-summer return level of number of elderly deaths in the period 2010 to 2030 using model TO and bias-corrected temperatures (BC-LSB for London; BC-LS for Budapest), with (a) elderly population assumed to be constant from mid-2003 in London and from mid-2001 in Budapest and (b) projected elderly population for each location.

Compared to London, the elderly population in Budapest is projected to increase at a slower rate [refer to Fig. 3.3(b)]. Using these population projections, the ensemble mean 2-summer mortality level in the period 2010 to 2030 is projected to be 12% lower than the present-day 2-summer level. This means that despite the projected increase in population and extreme temperatures in Budapest, the extreme daily number of elderly deaths in this period is projected to remain below the level observed in the 1990s.

6.5 Summary

This chapter has used mortality models and calibrated HadRM3 temperatures from previous chapters to project the future changes in extreme mortalities in London and Budapest. If elderly population sizes are assumed to be unchanged, the estimated change in the 2-

summer return level of number of elderly deaths relative to the present-day ranges from -12% to $+75\%$ for London in the period 2070 to 2099, across different HadRM3 perturbed physics ensemble members and different choices of mortality model and temperature calibration method. The corresponding change for Budapest ranges from -16% to $+22\%$. Uncertainties in climate model parameters and the choice of temperature calibration method have comparable effects on the uncertainties in extreme mortalities at both locations. The choice of mortality model has a smaller contribution to the uncertainties in extreme mortalities for London, while for Budapest, extreme mortality projections are not sensitive to the choice of mortality model.

Chapter 7

Conclusions and future work

7.1 Summary of main findings

This thesis has developed techniques for the projection of extreme heat-related elderly mortalities in London and Budapest from 2010 to 2099, using temperature projections from the perturbed physics ensemble (PPE) of the regional climate model (RCM) HadRM3. The general approach adopted here to project future mortalities is similar to that used previously: A statistical mortality model is developed to model the present-day relationship between daily mortality counts and observed weather variables. Future projections of relevant weather variables from climate models are calibrated to account for any discrepancies in model simulations. Future mortalities are then projected by driving the mortality model with calibrated climate model projections.

Despite the similarity in the general approach, this thesis is different from previous studies in several aspects. The primary focus is on the projection of future extreme daily heat-related mortality counts instead of the total number of heat-related deaths. For the climate model which provides future climate projections, this thesis has used the RCM HadRM3 with horizontal grid spacing of about 25 km, which is a finer resolution compared to the models used previously (e.g. 50 km in Dessai, 2002, 2003). Projections from the PPE of HadRM3 are also used to evaluate how mortality projections are affected by climate model uncertainty, a source of uncertainty not considered in previous studies.

This thesis has also investigated mortality modelling and temperature calibration methodologies. These are described below as the main findings at each stage of the mortality projection work are summarized.

Poisson generalized additive models have been used to model the present-day relationship between daily number of elderly deaths and weather observations. Changes in elderly population and non-weather related annual cycles in mortalities are modelled using an offset term and a non-parametric smooth function respectively. A number of candidate models have been considered. These models have different choices of environmental covariates, for example using combinations of temperature and humidity and using a biometeorological index. The functional dependence of daily mortality counts on the covariates is also specified in different ways. Model evaluation in previous studies was limited. In this thesis, a number of criteria have been used to evaluate the adequacy and specification of these candidate models. In particular, the mean residual mortality plots have assessed the ability of the models to predict the upper tails of present-day mortality distributions.

Considering these criteria, since the inclusion of humidity or a biometeorological index does not greatly improve model predictions, gridded observed daily mean air temperature has been chosen to be the only weather variable in the model for mortality projections. For London, between the model which specifies the mortality-temperature relationship using a non-parametric smooth function and the model which specifies such a relationship using a second order polynomial, there are large differences in the predicted mortality counts at extremely high temperatures. The former model predicts the upper tail of the mortality distribution better, but it is less robust to outliers which are related to the large daily number of deaths observed during the heatwave in 2003. On the other hand, the predictions for Budapest by the two models are similar.

The calibration of HadRM3 temperature projections is a major focus of this thesis. Previous mortality projections and other climate change impact studies typically adopted simplistic untested calibration methods. This thesis has considered two different approaches, bias correction and change factor, to calibrate the entire distribution of HadRM3 projected temperatures. For the bias correction approach, based on assumptions on the relationships in the location, scale and shape between the distributions of present-day observed temper-

atures and HadRM3 simulated temperatures, the transfer functions of three methods have been developed. In particular, a novel method has been proposed to correct the bias in the shape of the distribution of simulated temperatures. This involves first applying the Box-Cox transformation on these temperature distributions, then correcting the biases in the location and scale of the transformed distributions. The calibration of present-day HadRM3 simulated temperatures over Europe has revealed that this method gives the best result in correcting the biases in the upper tail of HadRM3 temperatures, compared to the other two methods which assumed no biases in shapes. Two change factor methods have also been considered. Their transfer functions have been developed based on assumptions on the relationships in the location and scale between the distributions of present-day and future HadRM3 temperatures.

The impacts of using different calibration approaches on extreme temperature projections have been investigated for Europe. The preferred methods for each approach, bias correction in location and scale of Box-Cox transformed temperatures and change factor in location and scale, give substantially different results. There is considerable difference in the spatial distribution of warming of extreme temperatures calibrated by the two approaches. In terms of the projected 10-summer temperature return level, the differences between the two approaches are greater than 4 °C over many parts of Europe in the period 2070 to 2099. Nevertheless, the 10-summer levels of temperatures calibrated by both methods are much lower (by more than 10 °C in certain places) than that without calibration. In addition, projections from the 11-member PPE of HadRM3 for London and Budapest have been calibrated. The results have suggested that uncertainty in climate model parameters and the uncertainty in the choice of calibration approach have comparable effects on extreme temperature projections.

Finally, using the mortality models and calibrated temperature projections, future mortalities have been projected for London and Budapest. These are followed by an assessment of changes in extreme daily mortality counts. Mortality projections are subject to uncertainties in mortality modelling, climate projections and the population affected by extreme heat. The uncertainty analyses in this thesis have focused on the sensitivity of extreme mortality projections to the choice of mortality model (parametric or non-parametric

specification of mortality-temperature relationship), climate model parameter uncertainty (calibrated projections from the PPE) and the choice of temperature calibration approach (bias correction or change factor). While these do not constitute a complete uncertainty assessment, these three sources of uncertainties have not been considered by previous mortality projections.

Incorporating the above-mentioned uncertainties, the possible range of future changes in extreme mortalities is large. If the elderly population size is assumed to remain constant in the future, the projected change in the 2-summer return level of number of daily elderly deaths for London ranges from -12% to $+75\%$ in the period 2070 to 2099 relative to the the present-day. The corresponding projected change for Budapest ranges from -16% to $+22\%$. Similar to extreme temperatures, uncertainties due to the choice of climate model parameters and the choice of temperature calibration method have comparable effects on the overall uncertainties in extreme mortality projections. The effect of the choice of mortality model has been found to be relatively less important.

7.2 Wider applications

Apart from human health, extreme heat has wide-ranging potential impacts on the society. Examples include reduced crop yield (Easterling et al., 2007), damage on transport infrastructure such as roads and rail tracks (McGregor et al., 2007) and power outages due to increased electricity demand for cooling (Miller et al., 2008). While this thesis has focused on the projection of extreme heat-related mortalities, many of the results obtained and techniques developed can be applied to the impact assessment of extreme heat and even other climate extremes.

The relationship between other climate change impacts and weather variables can be modelled in a similar way as how heat-related mortalities have been modelled in Chapter 4. The model evaluation procedure used in this thesis, especially the assessment of model predictions at extreme temperatures, can also be adopted when modelling such impacts.

With suitable impact models, extreme calibrated temperatures in Europe presented in

Chapter 5 can give useful indications of how other impacts of extreme heat will change in the future. In addition, the general calibration framework can be applied to climate model projections of other climate variables, for example rainfall. However, modifications might have to be made based on the characteristics of the distributions of the observed and model simulated climate variable of interest.

7.3 Future research directions

The assessment of heat-related health impacts of climate change covers multiple areas of statistical work, including regression analysis of the relationship between weather and health and the calibration of climate model projections. While this thesis has made improvements over previous mortality projections, there are certainly many aspects where further progress can be made in future research.

For the modelling of heat-related mortalities, the number of deaths observed during the 2003 heatwave in London has been found to be considerably higher than that predicted by mortality models used in this thesis. Recent research in epidemiology (e.g. Hajat et al., 2006) suggests a possible ‘heatwave effect’ where the number of deaths increases further in continuous periods of hot weather. It is worthwhile to consider how this effect can be incorporated into the heat-related mortality models, for example through the use of random effects mixed models (Chapter 6 in Wood, 2006).

As for the calibration of climate model projections, the techniques adopted in this thesis have been applied on the entire distribution of model projected temperatures. Given the importance of projecting the extremes of climate variables in impact studies, it would be interesting to develop alternative techniques which focus on the calibration of the tails of distributions of projected climate variables. Calibration involving multiple variables can also be considered. For example, it has been suggested that the bias in HadRM3 extreme temperatures in London is related to excessive soil drying in the model. Future work can attempt to incorporate the observed and modelled soil moisture variable in the calibration of temperature projections. Furthermore, model projected temperatures have been calibrated for individual time-slices. Improvements to the calibration procedure can

be explored such that the non-stationarity in the climate variable, e.g. long-term trends and annual cycles, can be accounted for.

Finally, the uncertainties of mortality projections have been assessed by analysing the sensitivity of projections to individual sources of uncertainties. A major area of current climate research has been to develop probabilistic projections of climate change based on model ensembles (e.g. PPE used in this thesis) constrained by observations, such that the combined effects of different sources of uncertainties can be quantified (Collins, 2007 and references therein; Murphy et al., 2009). Future research on mortality projections and other climate change impacts can consider how such ideas can be applied to assess uncertainties systematically.

Appendix A

Population projections

In the projection of future heat-related mortalities in London and Budapest, two sets of elderly population projection data are used (UK Office of National Statistics, 2008c; Eurostat, 2009). The methodologies used and the assumptions involved, such as the projected rates of mortality, regional and international migration, can be found in UK Office of National Statistics (2008a). Projections of elderly population of London are included as part of the 2006-based subnational UK population projection, and are therefore directly used in this study. The projected elderly population in Budapest, however, is derived from the 2004-based European Union regional population projection data up to 2031 (Eurostat, 2009), obtained from the Statistical Office of the European Communities, Eurostat. This appendix outlines the derivation methods.

The seven statistical regions of Hungary defined in the Eurostat data are shown in Fig. A.1. The city Budapest is located within the Közép-Magyarország region. The estimated elderly population sizes for Budapest and the Közép-Magyarország region from 1992 to 2007 (Hungarian Central Statistical Office, 2009) are shown by solid lines in Fig. A.2. The changes in elderly population sizes at the two locations had different signs: a decrease of about 0.5% per year in Budapest (except between 2000 and 2001) but an increase of about 0.6% per year in the Közép-Magyarország region. This means that the proportion of elderly population living in Budapest decreased steadily, as shown in Fig. A.3. A possible reason for this phenomenon is the migration of population out of the city area.



Figure A.1: Map showing the seven statistical regions of Hungary and the location of Budapest. Source: Hungarian Central Statistical Office, available at http://portal.ksh.hu/portal/page?_pageid=38,566887&_dad=portal&_schema=PORTAL.

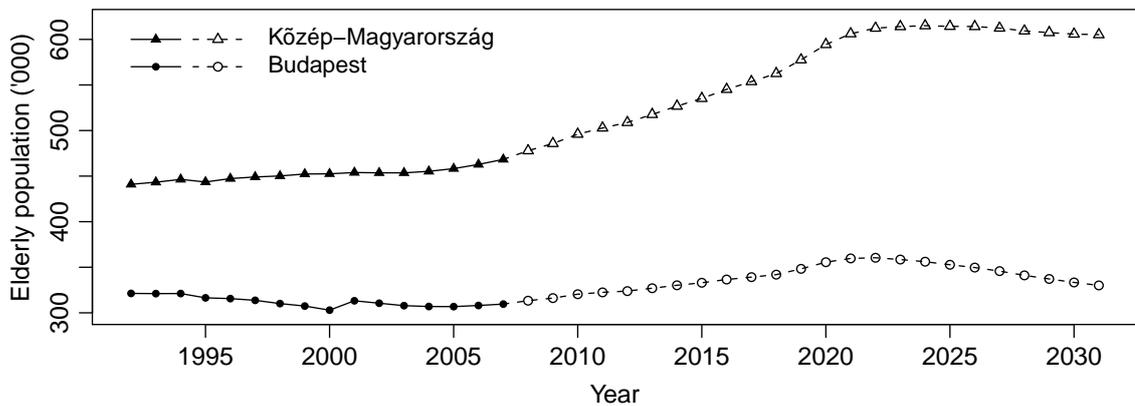


Figure A.2: Observed and projected elderly population sizes in Budapest (dots) and the Közép-Magyarország region of Hungary (triangles). The estimated population sizes from 1992 to 2007 are shown in solid lines. The projected sizes from 2007 to 2031 are shown in dashed lines.

According to the European Union projections, the elderly population size in the Közép-Magyarország region will increase gradually before levelling off at around 2020 (dashed line with triangles in Fig. A.2). In deriving the elderly population projection of Budapest from the projection of the Közép-Magyarország region, the approximately linear decreasing trend in the proportion of elderly population living within Budapest is assumed to continue, as shown by the dashed line in Fig. A.3. As a result, the Budapest elderly pop-

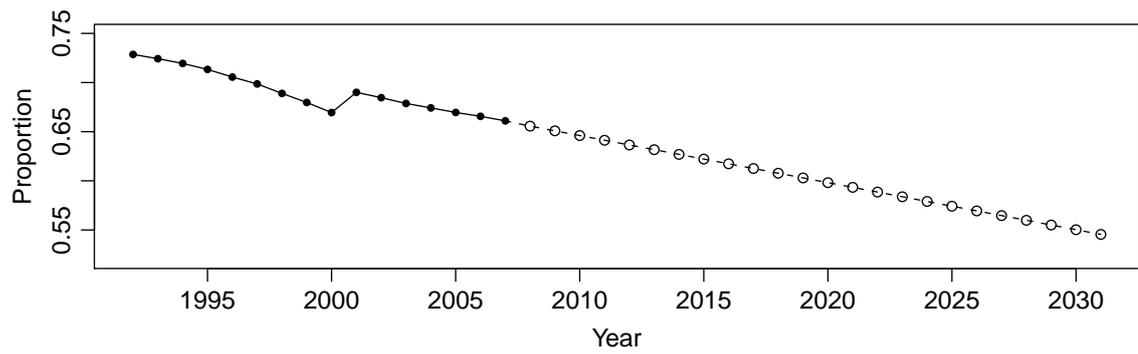


Figure A.3: Proportion of Budapest elderly population living in the Közép-Magyarország region of Hungary. The observed proportion from 1992 to 2007 are shown in solid line. The extrapolated proportion from 2007 to 2031 is shown in dashed line.

ulation size is projected to increase at a slower rate in the 2010s, then to fall afterwards (dashed line with circles in Fig. A.2).

Appendix B

Generalized linear models and generalized additive models

This appendix supplements Section 4.2.1 by providing further details on the estimation and evaluation of statistical mortality models. All of the mortality models considered in Section 4.4 are generalized additive models (GAMs). However, generalized linear models (GLMs), another type of models commonly used in modelling the relationship between mortality and weather, are first described because of the similarities between the two families.

B.1 Exponential family distribution

The random components of generalized linear models (GLMs) and generalized additive models (GAMs) assume the response variable to have a distribution from the exponential family, for example Poisson, Binomial, Gamma and Normal. Suppose there are n observations $\{y_1, y_2, \dots, y_n\}$ of the response variable Y_i . The probability density function (p.d.f.) of the exponential family can be written as

$$h(y_i; \theta_i, \phi) = \exp\left(\frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)\right) \quad (\text{B.1})$$

where $a(\cdot)$, $b(\cdot)$ and $c(\cdot)$ are known functions that define the type of distribution, and θ_i and ϕ represent the canonical parameter and dispersion parameter respectively. It can be

shown that the population mean and variance of the response variable Y are given by

$$\mathbb{E}(Y_i) = b'(\theta_i) \quad (\text{B.2})$$

$$\text{var}(Y_i) = b''(\theta_i)a(\phi). \quad (\text{B.3})$$

where b' and b'' represent the first and second derivative respectively.

In the case of a Poisson distribution (the usually assumed distribution for daily mortality counts) with a mean rate parameter μ whose p.d.f. is given by

$$h(y_i) = \frac{\mu_i^{y_i} \exp(-\mu_i)}{y_i!}. \quad (\text{B.4})$$

It can be shown by taking logs of (B.4) that for a Poisson distribution $\theta_i = \log(\mu_i)$, $\phi = 1$, $a = 1$, $b(\theta_i) = \exp(\theta_i)$ and $c(y_i, \phi) = -\log(y_i!)$ in (B.1). Substituting into (B.2) and (B.3), $\mathbb{E}(Y_i) = \text{var}(Y_i) = \mu_i$, i.e. the mean equals the variance.

B.2 Generalized linear models

As described in Section 4.2.1, the systematic component of a GLM has the structure

$$\eta_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} = \mathbf{X}_i \boldsymbol{\beta} \quad (\text{B.5})$$

where η_i is known as the linear predictor, $\mathbf{X}_i = \{X_{i1}, X_{i2}, \dots, X_{ip}\}$ is a set of p covariates, and $\boldsymbol{\beta} = \{\beta_1, \beta_2, \dots, \beta_p\}$ are unknown parameters to be estimated. The random and systematic components of the GLM are linked by a link function $g(\mu_i)$:

$$g(\mu_i) = \eta_i \quad (\text{B.6})$$

where $\mu_i = \mathbb{E}(Y_i)$. For a Poisson response, a log link function $g(\mu_i) = \log(\mu_i)$ is commonly used.

Rewriting $a_i(\phi) = \phi/\omega_i$, where ω_i are a set of known constants (often assumed to be

unity), the log-likelihood of β is given by

$$\begin{aligned}
l(\beta) &= \sum_{i=1}^n \log [f_{\theta_i}(y_i)] \\
&= \sum_{i=1}^n \frac{y_i \theta_i - b_i(\theta_i)}{a_i(\phi)} + c_i(\phi, y_i) \\
&= \sum_{i=1}^n \frac{\omega_i [y_i \theta_i - b_i(\theta_i)]}{\phi} + c_i(\phi, y_i). \tag{B.7}
\end{aligned}$$

It can be seen from (B.2), (B.3) and (B.6) how θ_i is linked to the model parameters β . By maximizing (B.7), β can be estimated. This is done using the iteratively re-weighted least squares (Nelder and Wedderburn, 1972). Under large sample approximation, the vector of predicted model parameter $\hat{\beta} \sim N(\beta, (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1})$, i.e. normally distributed with mean β and variance $(\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1}$, where \mathbf{W} is a weighting matrix. This can be used to construct confidence intervals of $\hat{\beta}$ and $\hat{\mu}$.

A measure of the goodness of fit of the model is the quantity *deviance*. It is analogous to the residual sum of squares in normal linear models, and is defined as

$$\begin{aligned}
D(\hat{\beta}) &= 2 \left[l(\hat{\beta}_{\max}) - l(\hat{\beta}) \right] \phi \\
&= \sum_{i=1}^n 2\omega_i \left[y_i (\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i) \right] \\
&= \sum_{i=1}^n d_i
\end{aligned}$$

where $l(\hat{\beta}_{\max})$ is the likelihood of the *saturated model* ($p = n$) which has the same number of parameters as the number of observations. It is thus the highest possible value of likelihood given the data. $\hat{\theta}_i$ and $\tilde{\theta}_i$ denote the maximum likelihood estimates of the canonical parameters for the model of interest and the saturated model, respectively. On the last line the contribution of deviance by the i^{th} observation is denoted by d_i .

As in normal linear models, the examination of residuals is an important part of the model checking procedure. However, the use of raw residuals is not appropriate because their variance is generally not constant. *Deviance residuals*, ϵ_i^d (Pierce and Schafer, 1986), defined as below, are used instead:

$$\hat{\epsilon}_i^d = \text{sign}(y_i - \hat{\mu}_i) \sqrt{d_i}$$

where $\hat{\mu}_i$ is the model predicted mean response. For a well-specified model, under sufficiently large samples, $\epsilon^d \sim N(0, 1)$.

B.3 Generalized additive models

Generalized additive models (GAMs) can be considered as an extension of the GLM framework. They have the structure

$$g(\mu_i) = \mathbf{X}_i\boldsymbol{\beta} + \sum_{k=1}^p f_k(X_{ik}) \quad (\text{B.8})$$

where the first term of the right hand side is a linear parametric model component, and f_j ($j = 1, 2, \dots, k$) represents non-parametric smooth functions of covariates. The combined effects of multiple explanatory variables (interaction) can be included by adding functions of multiple covariates [e.g. $f(X_1, X_2)$].

In order to estimate the model, each smooth function f_j is specified as a linear combination of basis functions b_{jk} such that

$$f_j(X_j) = \sum_{k=1}^q \beta_{jk}^* b_{jk}(X_j)$$

where β_{jk}^* are the coefficients of the smoothing functions. In this way, (B.8) can then be written as $g(\mu_i) = \mathbf{X}_i\boldsymbol{\beta}$, same as (B.5).

A number of possible functional bases, such as polynomial or cubic splines, can be chosen according to the application (e.g. some bases cannot be used to smooth w.r.t. more than one covariate) and computational efficiency. Cubic regression splines are used in fitting the GAMs in Section 4.4. These are knot-based splines where the knots are spread evenly across the range the covariate values. For each smoothing function, a basis dimension is specified. This can be considered as the upper limit of the degrees of freedom (d.f.) of the smoothing function, but the actual (effective) d.f. are controlled by penalization of over-smoothing described below. Details of different types of smoothing basis functions can be found in Chapter 4.1 of Wood (2006).

For GAMs, the parameters are not estimated by maximizing the log-likelihood (B.7) because this will result in over-fitting by having over-wiggly smoothing functions. In-

stead a penalty term $\sum_j \lambda_j \int [f_j''(x)]^2 dx$ is added to the log likelihood, where λ_j represents smoothing parameters to be estimated as described below. Each penalty term can be written in terms of the quadratic form in terms of $\boldsymbol{\beta}$ (Chapter 4.2 in Wood, 2006):

$$\int [f''(x)]^2 dx = \boldsymbol{\beta}^T \mathbf{S} \boldsymbol{\beta},$$

where the matrix \mathbf{S} contains known coefficients. Therefore the penalized log-likelihood $l_p(\boldsymbol{\beta})$ to be maximized is defined as

$$l_p(\boldsymbol{\beta}) = l(\boldsymbol{\beta}) - \frac{1}{2} \sum_j \lambda_j \boldsymbol{\beta}^T \mathbf{S}_j \boldsymbol{\beta}$$

which represents a trade off between the goodness of fit (first term on the right hand side) and model smoothness (second term on the right hand side). The maximum likelihood estimates of the parameters are obtained by iteration to convergence by a process known as penalized iteratively re-weighted least squares.

A number of approaches can be used to estimate the smoothing parameters λ_j (see Chapter 4 in Wahba, 1990). For the case of Poisson response where the scale parameter is known, this is done by minimizing the un-biased risk estimator (UBRE), a deviance-based measure:

$$\text{UBRE} = \frac{1}{n_o} D(\hat{\boldsymbol{\beta}}) + \frac{2n_p}{n_o}$$

where n_p is the number of degrees of freedom in the model, and n_o is the number of observations. The UBRE score is also a measure that can be used to assess relatively the predictive power of models. A model with a lower UBRE score is preferred. As in GLMs, model checking for GAMs also involve the examination of deviance residuals as described in the last section.

Appendix C

Generalized Pareto distribution

In Chapters 4 and 5, Generalized Pareto (GP) distributions are fitted to excesses of observed and HadRM3 temperatures over thresholds in order to estimate the probabilities of extreme temperature occurrences. This appendix provides a brief overview of this distribution. Further details can be found in Coles (2001) and Beirlant et al. (2004). In the following discussion, consider a summer daily mean air temperature variable denoted by T .

Let $\{T_1, T_2, \dots, T_n, \dots, T_{ny}\}$ be the daily sequence of T over a period of y years with the length of summer defined to be n days. Let also M_n be the annual maximum of summer daily mean temperature, e.g. $M_n = \max\{T_1, T_2, \dots, T_n\}$. The daily temperature series values are assumed to be independent and identically distributed. If there exists sequences of constants $\{a_n\}$ and $\{b_n\}$ such that, as $n \rightarrow \infty$,

$$Pr\left(\frac{M_n - b_n}{a_n} \leq z\right) \rightarrow G(z),$$

then G takes the following form of generalized extreme value (GEV) distribution

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \quad (\text{C.1})$$

with location parameter μ , scale parameter $\sigma > 0$ and shape parameter ξ . See Kharin et al. (2007) for an example of how changes in the annual extremes of model projected climate variables are estimated using the GEV distribution.

Now consider a random variable $V = T - u$ to represent excesses of T above a threshold u . In the asymptotic limit of a large u ,

$$Pr(V \leq v | V > 0) = H(v) = 1 - \left(1 + \frac{\xi v}{\tilde{\sigma}}\right)^{-\frac{1}{\xi}}, \quad (\text{C.2})$$

where $v > 0$ and $(1 + \xi v/\tilde{\sigma}) > 0$. This is known as the generalized Pareto distribution. Here the scale parameter $\tilde{\sigma}$ depends on the threshold and is related to σ in the corresponding GEV distribution (C.1) by $\tilde{\sigma} = \sigma + \xi(u - \mu)$, while the shape parameter ξ is the same as that in the GEV distribution. If $\xi < 0$ the distribution of excesses has a upper limit of $u - \tilde{\sigma}/\xi$, whereas if $\xi > 0$ it has infinite upper limit. For $\xi = 0$, the GP distribution becomes the exponential distribution:

$$H(v) = 1 - \exp\left(-\frac{v}{\tilde{\sigma}}\right). \quad (\text{C.3})$$

The choice of threshold u involves a bias-variance tradeoff. With a lower threshold, more excesses can be used to estimate the model so that the variance of parameter estimates is reduced. However, a sufficiently high threshold is needed to ensure that the asymptotic basis of the model is satisfied. This selection of threshold is subjective and can be aided by diagnostics such as the ‘mean residual life plots’ and plots showing the stability of model parameters across different thresholds (Chapter 4 in Coles, 2001). Since this procedure is not practical for this thesis where the GP distribution needs to be fitted to large gridded data sets with many variables, we use a constant threshold $u = \hat{T}_{0.95}$, the sample 0.95 quantile.

The parameters of the GP distribution can be estimated by maximum likelihood. For this thesis this is done using the ‘evd’ package in R. From (C.2) and (C.3), the exceedance probability of T exceeding the level T_m for $T_m > u$ is given by

$$Pr(T > T_m) = \frac{1}{r_m} = \begin{cases} \zeta \left[1 + \xi \left(\frac{T_m - u}{\tilde{\sigma}}\right)\right]^{-\frac{1}{\xi}} & \text{for } \xi \neq 0 \\ \zeta \exp\left(-\frac{T_m - u}{\tilde{\sigma}}\right) & \text{for } \xi = 0, \end{cases} \quad (\text{C.4})$$

where $\zeta = Pr(T > u)$ which can be estimated by the observed proportion of threshold exceedances in the sample ($\zeta = 0.05$ for the choice $u = \hat{T}_{0.95}$). As T_m is the level which is exceeded on average every r_m days, T_m is known as the *return level* with a *return period*

of r_m days. Using (C.4), the return levels for different return periods can be estimated by the estimated GP distribution parameters:

$$\hat{T}_m = \begin{cases} u + \frac{\hat{\sigma}}{\hat{\xi}} \left[(r_m \hat{\zeta})^{\hat{\xi}} - 1 \right] & \text{for } \hat{\xi} \neq 0 \\ u + \hat{\sigma} \log(r_m \hat{\zeta}) & \text{for } \hat{\xi} = 0. \end{cases}$$

For example, with n days of observations in each summer, the 10-summer return level is found by setting $r_m = 10n$. A return level plot of \hat{T}_m against r_m (typically on a logarithmic scale to emphasize the extremes) can be used to evaluate the model fit. The empirical estimates of return levels can be added to the plot as a diagnosis of the model fit. For gridded data sets, spatial maps of return levels estimated by the GP distribution fit and spatial maps of sample quantiles with the same exceedance probability have been compared, but such comparisons are not shown in Chapter 5 for the sake of brevity.

Bibliography

- Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Kumar, K. R., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D. B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., and Vazquez-Aguirre, J. L. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research*, 111(D5):D05109, doi:10.1029/2005JD006290.
- Anders, I. and Rockel, B. (2009). The influence of prescribed soil type distribution on the representation of present climate in a regional climate model. *Climate Dynamics*, 33(2-3):177–186, doi:10.1007/s00382-008-0470-y.
- Anderson, B. G. and Bell, M. L. (2009). Weather-related mortality: How heat, cold, and heat waves affect mortality in the United States. *Epidemiology*, 20(2):205–213, doi:10.1097/ede.0b013e318190ee08.
- Anderson, H. R., Derwent, R. G., and Stedman, J. (2002). Overview of climate change impacts on human health in the UK: Air Pollution and Climate Change. In *Health effects of climate change in the UK*, chapter 4.7, pages 193–213. Department of Health, UK.
- Armstrong, B. (2006). Models for the relationship between ambient temperature and daily mortality. *Epidemiology*, 17(6):624–631, doi:10.1097/01.ede.0000239732.50999.8f.
- Baccini, M., Biggeri, A., Accetta, G., Kosatsky, T., Katsouyanni, K., Analitis, A., Anderson, H. R., Bisanti, L., D’Ippoliti, D., Danova, J., Forsberg, B., Medina, S., Paldy, A., Rabczenko, D., Schindler, C., and Michelozzi, P. (2008). Heat effects on mortality in 15 European cities. *Epidemiology*, 19(5):711–719, doi:10.1097/ede.0b013e318176bfcd.

- Ballester, F., Corella, D., Pérez-Hoyos, S., Sáez, M., and Hervás, A. (1997). Mortality as a function of temperature. A study in Valencia, Spain, 1991-1993. *International Journal of Epidemiology*, 26(3):551–561.
- Basu, R. and Samet, J. M. (2002). Relation between elevated ambient temperature and mortality: A review of the epidemiologic evidence. *Epidemiologic Reviews*, 24(2):190–202, doi:10.1093/epirev/mxf007.
- Beirlant, J., Goegbeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*. Wiley.
- Beniston, M. and Diaz, H. F. (2004). The 2003 heat wave as an example of summers in a greenhouse climate? Observations and climate model simulations for Basel, Switzerland. *Global and Planetary Change*, 44(1):73–81, doi:10.1016/j.gloplacha.2004.06.006.
- Beniston, M., Stephenson, D. B., Christensen, O. B., Ferro, C. A. T., Frei, C., Goyette, S., Halsnaes, K., Holt, T., Jylhä, K., Koffi, B., Palutikof, J., Schöll, R., Semmler, T., and Woth, K. (2007). Future extreme events in european climate: An exploration of regional climate model projections. *Climatic Change*, 81(Suppl. 1):71–95, doi:10.1007/s10584-006-9226-z.
- Black, E., Blackburn, M., Harrison, G., Hoskins, B., and Methven, J. (2004). Factors contributing to the summer 2003 European heatwave. *Weather*, 59(8):217–223, doi:10.1256/wea.74.04.
- Blake, E. S., Rappaport, E. N., and Landsea, C. W. (2007). The deadliest, costliest, and most intense United States tropical cyclones from 1851 to 2006 (and other frequently related hurricane facts). *NOAA Technical Memorandum NWS TPC-5*, National Hurricane Center, National Weather Service.
- Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society Series B - Statistical Methodology*, 26(2):211–252.

- Braga, A. L. F., Zanobetti, A., and Schwartz, J. (2001). The time course of weather-related deaths. *Epidemiology*, 12(6):662–667.
- Braga, A. L. F., Zanobetti, A., and Schwartz, J. (2002). The effect of weather on respiratory and cardiovascular deaths in 12 U.S. cities. *Environmental Health Perspectives*, 110(9):859–863.
- Buonomo, E., Jones, R., Huntingford, C., and Hannaford, J. (2007). On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations. *Quarterly Journal of the Royal Meteorological Society*, 133(622):65–81, doi:10.1002/qj.13.
- Carril, A., Gualdi, S., Cherchi, A., and Navarra, A. (2008). Heatwaves in Europe: areas of homogeneous variability and links with the regional to large-scale atmospheric and SSTs anomalies. *Climate Dynamics*, 30(1):77–98, doi:10.1007/s00382-007-0274-5.
- Carson, C., Hajat, S., Armstrong, B., and Wilkinson, P. (2006). Declining vulnerability to temperature-related mortality in London over the 20th century. *American Journal of Epidemiology*, 164(1):77–84, doi:10.1093/aje/kwj147.
- Cassou, C., Terray, L., and Phillips, A. S. (2005). Tropical Atlantic influence on European heat waves. *Journal of Climate*, 18(15):2805–2811, doi:10.1175/JCLI3506.1.
- Centers for Disease Control and Prevention (2006). Extreme heat: A prevention guide to promote your personal health and safety. Available at http://www.bt.cdc.gov/disasters/extremeheat/heat_guide.asp. Accessed 1 April 2010.
- Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P. (2008). On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters*, 35(20):L20709, doi:10.1029/2008GL035694.
- Christensen, J. H., Carter, T. R., Rummukainen, M., and Amanatidis, G. (2007a). Evaluating the performance and utility of regional climate models: The PRUDENCE project. *Climatic Change*, 81(Suppl. 1):1–6, doi:10.1007/s10584-006-9211-6.

- Christensen, J. H. and Christensen, O. B. (2007). A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Climatic Change*, 81(Suppl. 1):7–30, doi:10.1007/s10584-006-9210-7.
- Christensen, J. H., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, I., Jones, R., Kolli, R. K., Kwon, W., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C. G., Räisänen, J., Rinke, A., Sarr, A., and Whetton, P. (2007b). Regional Climate Projections. In Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., editors, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 11, pages 847–940. Cambridge University Press.
- Clark, R. T., Brown, S. J., and Murphy, J. M. (2006). Modeling northern hemisphere summer heat extreme changes and their uncertainties using a physics ensemble of climate sensitivity experiments. *Journal of Climate*, 19(17):4418–4435, doi:10.1175/JCLI3877.1.
- Clayton, D. and Hills, M. (1998). *Statistical methods in epidemiology*. Oxford University Press.
- Coelho, C. A. S., Ferro, C. A. T., Stephenson, D. B., and Steinskog, D. J. (2007). Methods for exploring spatial and temporal variability of extreme events in climate data. *Journal of Climate*, 21:2072–2092, doi:10.1175/2007JCLI1781.1.
- Cohen, J., Veysseire, J., and Bessemoulin, P. (2005). Bio-climatological aspects of summer 2003 over France. In Kirch, W., Menne, B., and Bertollini, R., editors, *Extreme weather events and public health response*, chapter 4, pages 33–45. Springer.
- Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer.
- Collins, M. (2007). Ensembles and probabilities: a new era in the prediction of climate change. *Philosophical Transactions of the Royal Society A - Mathematical, Physical & Engineering Sciences*, 365(1857):1957–1970, doi:10.1098/rsta.2007.2068.

- Collins, M., Booth, B. B. B., Harris, G. R., Murphy, J. M., Sexton, D. M. H., and Webb, M. J. (2006). Towards quantifying uncertainty in transient climate change. *Climate Dynamics*, 27(2-3):127–147, doi:10.1007/s00382-006-0121-0.
- Confalonieri, U., Menne, B., Akhtar, R., Ebi, K. L., Hauengue, M., Kovats, R. S., Revich, B., and Woodward, A. (2007). Human health. In Parry, M. L., Canziani, O. F., Palutikof, J. L., van der Linden, P. J., and Hanson, C. E., editors, *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 8, pages 391–431. Cambridge University Press.
- Cressie, N. (1993). *Statistics for Spatial Data*. Wiley.
- Curriero, F. C., Heiner, K. S., Samet, J. M., Zeger, S. L., Strug, L., and Patz, J. A. (2002). Temperature and mortality in 11 cities of the eastern United States. *American Journal of Epidemiology*, 155(1):80–87.
- D’Andrea, F., Tibaldi, S., Blackburn, M., Boer, G., Déqué, M., Dix, M. R., Dugas, B., Ferranti, L., Iwasaki, T., Kitoh, A., Pope, V., Randall, D., Roeckner, E., Straus, D., Stern, W., Van den Dool, H., and Williamson, D. (1998). Northern Hemisphere atmospheric blocking as simulated by 15 atmospheric general circulation models in the period 1979-1988. *Climate Dynamics*, 14(6):385–407, doi:10.1007/s003820050230.
- Davis, R. E., Knappenberger, P. C., Michaels, P. J., and Novicoff, W. M. (2003). Changing heat-related mortality in the United States. *Environmental Health Perspectives*, 111(14):1712–1718, doi:10.1289/ehp.6336.
- Davis, R. E., Knappenberger, P. C., Michaels, P. J., and Novicoff, W. M. (2004). Seasonality of climate-human mortality relationships in US cities and impacts of climate change. *Climate Research*, 26(1):61–76.
- Davison, A. C. (2003). *Statistical Models*. Cambridge University Press.
- Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and Their Applications*. Cambridge University Press.

- de' Donato, F. K., Stafoggia, M., Rognoni, M., Poncino, S., Caranci, N., Bisanti, L., Demaria, M., Forastiere, F., Michelozzi, P., Pelosini, R., and Perucci, C. A. (2008). Airport and city-centre temperatures in the evaluation of the association between heat and mortality. *International Journal of Biometeorology*, 52(4):301–310, doi:10.1007/s00484-007-0124-5.
- Della-Marta, P. M., Luterbacher, J., von Weissenfluh, H., Xoplaki, E., Brunet, M., and Wanner, H. (2007). Summer heat waves over western Europe 1880-2003, their relationship to large-scale forcings and predictability. *Climate Dynamics*, 29(2-3):251–275, doi:10.1007/s00382-007-0233-1.
- Déqué, M., Rowell, D. P., Lüthi, D., Giorgi, F., Christensen, J. H., Rockel, B., Jacob, D., Kjellström, E., de Castro, M., and van den Hurk, B. (2007). An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Climatic Change*, 81(Suppl. 1):53–70, doi:10.1007/s10584-006-9228-x.
- Dessai, S. (2002). Heat stress and mortality in Lisbon Part I. model construction and validation. *International Journal of Biometeorology*, 47(1):6–12, doi:10.1007/s00484-002-0143-1.
- Dessai, S. (2003). Heat stress and mortality in Lisbon Part II. An assessment of the potential impacts of climate change. *International Journal of Biometeorology*, 48(1):37–44, doi:10.1007/s00484-003-0180-4.
- Donaldson, G., Kovats, R. S., Keatinge, W. R., and McMichael, A. J. (2002). Overview of climate change impacts on human health in the UK: Heat- and cold-related mortality and morbidity and climate change. In *Health effects of climate change in the UK*, chapter 4.1, pages 70–79. Department of Health, UK.
- Donaldson, G. C., Keatinge, W. R., and Nayha, S. (2003). Changes in summer temperature and heat-related mortality since 1971 in North Carolina, South Finland, and Southeast England. *Environmental Research*, 91(1):1–7.

- Draper, N. R. and Smith, H. (1998). *Applied Regression Analysis*. John Wiley & Sons, New York, USA, 3rd edition.
- Easterling, W. E., Aggarwal, P. K., Batima, P., Brander, K. M., Erda, L., Howden, S. M., Kirilenko, A., Morton, J., Soussana, J.-F., Schmidhuber, J., and Tubiello, F. N. (2007). Food, fibre and forest products. In Parry, M. L., Canziani, O. F., Palutikof, J. L., van der Linden, P. J., and Hanson, C. E., editors, *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 5, pages 273–313. Cambridge University Press.
- Eastoe, E. F. and Tawn, J. A. (2009). Modelling non-stationary extremes with application to surface level ozone. *Journal of the Royal Statistical Society Series C - Applied Statistics*, 58(Part 1):25–45, doi:10.1111/j.1467-9876.2008.00638.x.
- Ebi, K. L. (2008). Healthy people 2100: modeling population health impacts of climate change. *Climatic Change*, 88(1):5–19, doi:10.1007/s10584-006-9233-0.
- Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer-Verlag.
- Eurostat (2008). *Eurostat regional yearbook 2008*. European Communities.
- Eurostat (2009). EUROPOP2004 - Trend scenario, regional level. Available at <http://epp.eurostat.ec.europa.eu/portal/page/portal/population/data/database>. Accessed 1 April 2010.
- Fanger, P. O. (1970). *Thermal comfort: Analysis and applications in environmental engineering*. Krieger, New York.
- Ferro, C. A. T., Hannachi, A., and Stephenson, D. B. (2005). Simple nonparametric techniques for exploring changing probability distributions of weather. *Journal of Climate*, 18(21):4344–4354, doi:10.1175/JCLI3518.1.

- Fink, A. H., Brücher, T., Krüger, A., Leckebusch, G. C., Pinto, J. G., and Ulbrich, U. (2004). The 2003 European summer heatwaves and drought - synoptic diagnosis and impacts. *Weather*, 59(8):209–216, doi:10.1256/wea.73.04.
- Fischer, E. M. and Schär, C. (2009). Future changes in daily summer temperature variability: driving processes and role for temperature extremes. *Climate Dynamics*, 33(7-8):917–935, doi:10.1007/s00382-008-0473-8.
- Fischer, E. M., Seneviratne, S. I., Lüthi, D., and Schär, C. (2007a). Contribution of land-atmosphere coupling to recent European summer heat waves. *Geophysical Research Letters*, 34(6):L06707, doi:10.1029/2006GL029068.
- Fischer, E. M., Seneviratne, S. I., Vidale, P. L., Lüthi, D., and Schär, C. (2007b). Soil moisture-atmosphere interactions during the 2003 European summer heat wave. *Journal of Climate*, 20(20):5081–5099, doi:10.1175/JCLI4288.1.
- Folland, C. and Anderson, C. (2002). Estimating changing extremes using empirical ranking methods. *Journal of Climate*, 15(20):2954–2960, doi:10.1175/1520-0442(2002)015<2954:ECEUER>2.0.CO;2.
- Fouillet, A., Rey, G., Jouglu, E., Frayssinet, P., Bessemoulin, P., and Hémon, D. (2007). A predictive model relating daily fluctuations in summer temperatures and mortality rates. *BMC Public Health*, 7, doi:10.1186/1471-2458-7-114.
- Fouillet, A., Rey, G., Wagner, V., Laaidi, K., Empereur-Bissonnet, P., Le Tertre, A., Frayssinet, P., Bessemoulin, P., Laurent, F., De Crouy-Chanel, P., Jouglu, E., and Hémon, D. (2008). Has the impact of heat waves on mortality changed in France since the European heat wave of summer 2003? A study of the 2006 heat wave. *International Journal of Epidemiology*, 37(2):309–317, doi:10.1093/ije/dym253.
- Frich, P., Alexander, L. V., Della-Marta, P., Gleason, B., Haylock, M., Klein Tank, A. M. G., and Peterson, T. (2002). Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research*, 19(3):193–212, doi:10.3354/cr019193.

- Füssel, H., Klein, R. J. T., and Ebi, K. L. (2006). Adaptation assessment for public health. In Menne, B. and Ebi, K. L., editors, *Climate change and adaptation strategies for human health*, chapter 3, pages 41–62. Springer.
- Garcia-Herrera, R., Díaz, J., Trigo, R. M., and Hernández, E. (2005). Extreme summer temperatures in Iberia: health impacts and associated synoptic conditions. *Annales Geophysicae*, 23(2):239–251.
- Garssen, J., Harmsen, C., and de Beer, J. (2005). The effect of the summer 2003 heat wave on mortality in the Netherlands. *Eurosurveillance*, 10: Available at <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=557>.
- Giorgi, F., Hewitson, B., Christensen, J., Hulme, M., von Storch, H., Whetton, P., Jones, R., Mearns, L., and Fu, C. (2001). Regional Climate Information - Evaluation and Projections. In Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., and Johnson, C. A., editors, *Climate Change 2001: The Physical Science Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 10, pages 583–638. Cambridge University Press.
- Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B., and Wood, R. A. (2000). The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, 16(2-3):147–168, doi:10.1007/s003820050010.
- Gosling, S. N., Lowe, J. A., McGregor, G. R., Pelling, M., and Malamud, B. D. (2009a). Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. *Climatic Change*, 92(3-4):299–341, doi:10.1007/s10584-008-9441-x.
- Gosling, S. N., McGregor, G. R., and Lowe, J. A. (2009b). Climate change and heat-related mortality in six cities Part 2: climate model evaluation and projected impacts from changes in the mean and variability of temperature with climate change. *International Journal of Biometeorology*, 53(1):31–51, doi:10.1007/s00484-008-0189-9.

- Gosling, S. N., McGregor, G. R., and Paldy, A. (2007). Climate change and heat-related mortality in six cities Part 1: model construction and validation. *International Journal of Biometeorology*, 51(6):525–540, doi:10.1007/s00484-007-0092-9.
- Gouveia, N., Hajat, S., and Armstrong, B. (2003). Socioeconomic differentials in the temperature-mortality relationship in São Paulo, Brazil. *International Journal of Epidemiology*, 32(3):390–397, doi:10.1093/ije/dyg077.
- Grazzini, F., Ferranti, L., Lalaurette, F., and Vitart, F. (2003). The exceptional warm anomalies of summer 2003. *ECMWF Newsletter*, No. 99:30–31.
- Gregory, J. M., Mitchell, J. F. B., and Brady, A. J. (1997). Summer drought in northern midlatitudes in a time-dependent CO₂ climate experiment. *Journal of Climate*, 10(4):662–686, doi:10.1175/1520-0442(1997)010<0662:SDINMI>2.0.CO;2.
- Grize, L., Huss, A., Thommen, O., Schindler, C., and Braun-Fahrländer, C. (2005). Heat wave 2003 and mortality in Switzerland. *Swiss Medical Weekly*, 135:200–205.
- Guest, C. S., Willson, K., Woodward, A. J., Hennessy, K., Kalkstein, L. S., Skinner, C., and McMichael, A. J. (1999). Climate and mortality in Australia: retrospective study, 1979–1990, and predicted impacts in five major cities in 2030. *Climate Research*, 13(1):1–15, doi:10.3354/cr013001.
- Hacker, J. N., Belcher, S. E., and Connell, R. K. (2005). *Beating the Heat: keeping UK buildings cool in a warming climate*. UKCIP Briefing Report. UKCIP, Oxford.
- Haines, A., Kovats, R. S., Campbell-Lendrum, D., and Corvalan, C. (2006). Climate change and human health: Impacts, vulnerability and public health. *Public Health*, 120(7):585–596, doi:10.1016/j.puhe.2006.01.002. World Climate Change Conference, Moscow, Russia, Sep 29-Oct 03, 2003.
- Hajat, S., Armstrong, B., Baccini, M., Biggeri, A., Bisanti, L., Russo, A., Paldy, A., Menne, B., and Kosatsky, T. (2006). Impact of high temperatures on mortality - Is there an added heat wave effect? *Epidemiology*, 17(6):632–638, doi:10.1097/01.ede.0000239688.70829.63.

- Hajat, S., Armstrong, B. G., Gouveia, N., and Wilkinson, P. (2005). Mortality displacement of heat-related deaths - A comparison of Delhi, São Paulo, and London. *Epidemiology*, 16(5):613–620, doi:10.1097/01.ede.0000164559.41092.2a.
- Hajat, S., Kovats, R. S., Atkinson, R. W., and Haines, A. (2002). Impacts of hot temperatures on death in London: A time series approach. *Journal of Epidemiology and Community Health*, 56(5):367–372, doi:10.1136/jech.56.5.367.
- Hajat, S., Kovats, R. S., and Lachowycz, K. (2007). Heat-related and cold-related deaths in England and Wales: who is at risk? *Occupational and Environmental Medicine*, 64(2):93–100, doi:10.1136/oem.2006.029017.
- Hales, S., Edwards, S. J., and Kovats, R. S. (2003). Impacts on health of climate extremes. In McMichael, A. J., Campbell-Lendrum, D. H., Corvalán, C. F., Ebi, K. L., Githeko, A. K., Scheraga, J. D., and Woodward, A., editors, *Climate change and human health*, chapter 5, pages 79–102. World Health Organization.
- Harmsen, C. and Garssen, J. (2006). Statistics Netherlands: July heat causes one thousand extra deaths. Available at <http://www.cbs.nl/en-GB/menu/themas/bevolking/publicaties/artikelen/archief/2006/2006-2019-wm.htm>. Accessed 1 April 2010.
- Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized additive models*. Chapman & Hall.
- Havenith, G. (1999). Heat balance when wearing protective clothing. *Annals of Occupational Hygiene*, 43(5):289–296.
- Havenith, G. (2005). Temperature regulation, heat balance and climatic stress. In Kirch, W., Menne, B., and Bertollini, R., editors, *Extreme weather events and public health response*, chapter 7, pages 69–80. Springer.
- Hawkins, E. and Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, 90(8):1095–1107, doi:10.1175/2009BAMS2607.1.

- Hayhoe, K., Cayan, D., Field, C. B., Frumhoff, P. C., Maurer, E. P., Miller, N. L., Moser, S. C., Schneider, S. H., Cahill, K. N., Cleland, E. E., Dale, L., Drapek, R., Hanemann, R. M., Kalkstein, L. S., Lenihan, J., Lunch, C. K., Neilson, R. P., Sheridan, S. C., and Verville, J. H. (2004). Emissions pathways, climate change, and impacts on California. *Proceedings of the National Academy of Sciences of the United States of America*, 101(34):12422–12427, doi:10.1073/pnas.0404500101.
- Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M. (2008). A European daily high-resolution gridded dataset of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research*, 113:D20119, doi:10.1029/2008JD010201.
- Hegerl, G. C., Zwiers, F. W., Braconnot, P., Gillett, N. P., Luo, Y., Marengo Orsini, J. A., Nicholls, N., Penner, J. E., and Stott, P. A. (2007). Understanding and Attributing Climate Change. In Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., editors, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 9, pages 663–745. Cambridge University Press.
- Hofstra, N., New, M., and McSweeney, C. (2009). The influence of interpolation and station network density on the distribution and extreme trends of climate variables in gridded data. *Climate Dynamics*, doi:10.1007/s00382-009-0698-1. Published online.
- Hungarian Central Statistical Office (2009). Population statistics database. Available at <http://statinfo.ksh.hu/Statinfo/>. Accessed 1 April 2010.
- Huynen, M. M. T. E., Martens, P., Schram, D., Weijenberg, M. P., and Kunst, A. E. (2001). The impact of heat waves and cold spells on mortality rates in the Dutch population. *Environmental Health Perspectives*, 109(5):463–470.
- Hyndman, R. and Fan, Y. (1996). Sample quantiles in statistical packages. *American Statistician*, 50(4):361–365.

- Ines, A. V. M. and Hansen, J. W. (2006). Bias correction of daily GCM rainfall for crop simulation studies. *Agricultural and Forest Meteorology*, 138(1-4):44–53, doi:10.1016/j.agrformet.2006.03.009.
- Ishigami, A., Hajat, S., Kovats, R. S., Bisanti, L., Rognoni, M., Russo, A., and Paldy, A. (2008). An ecological time-series study of heat-related mortality in three European cities. *Environmental Health*, 7:5, doi:10.1186/1476-069X-7-5.
- Jacob, D., Bärring, L., Christensen, O. B., Christensen, J. H., de Castro, M., Déqué, M., Giorgi, F., Hagemann, S., Lenderink, G., Rockel, B., Sanchez, E., Schär, C., Seneviratne, S. I., Somot, S., van Ulden, A., and van den Hurk, B. (2007). An inter-comparison of regional climate models for Europe: model performance in present-day climate. *Climatic Change*, 81(Suppl. 1):31–52, doi:10.1007/s10584-006-9213-4.
- Johnson, H., Kovats, R. S., McGregor, G., Stedman, J., Gibbs, M., Walton, H., Cook, L., and Black, E. (2005). The impact of the 2003 heat wave on mortality and hospital admissions in England. *Health Statistics Quarterly*, 25:6–11.
- Jones, P. D., Kilsby, C. G., Harpham, C., Glenis, V., and Burton, A. (2009). *UK Climate Projections science report: Projections of future daily climate for the UK from the Weather Generator*. University of Newcastle, UK.
- Jones, P. D. and Lister, D. H. (2009). The urban heat island in Central London and urban-related warming trends in Central London since 1900. *Weather*, 64(12):323–327, doi:10.1002/wea.432.
- Jones, R. G., Murphy, J. M., and Noguer, M. (1995). Simulation of climate change over Europe using a nested regional-climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries. *Quarterly Journal of the Royal Meteorological Society*, 121(526):1413–1449, doi:10.1002/qj.49712152610.
- Jones, R. G., Noguer, M., Hassell, D. C., Hudson, D., Wilson, S. S., Jenkins, G. J., and Mitchell, J. F. B. (2004). *Generating high resolution climate change scenarios using PRECIS*. Met Office Hadley Centre, Exeter, UK.

- Kalkstein, L. S. and Davis, R. E. (1989). Weather and human mortality: an evaluation of demographic and interregional responses in the United States. *Annals of the Association of American Geographers*, 79(1):44–64, doi:10.1111/j.1467-8306.1989.tb00249.x.
- Kalkstein, L. S. and Greene, J. S. (1997). An evaluation of climate/mortality relationships in large U.S. cities and the possible impacts of a climate change. *Environmental Health Perspectives*, 105(1):84–93.
- Kan, H., London, S. J., Chen, H., Song, G., Chen, G., Jiang, L., Zhao, N., Zhang, Y., and Chen, B. (2007). Diurnal temperature range and daily mortality in Shanghai, China. *Environmental Research*, 103(3):424–431, doi:10.1016/j.envres.2006.11.009.
- Keatinge, W. R., Donaldson, G. C., Cordioli, E. A., Martinelli, M., Kunst, A. E., Mackenbach, J. P., Nayha, S., and Vuori, I. (2000). Heat related mortality in warm and cold regions of Europe: observational study. *British Medical Journal*, 321(7262):670–673.
- Kharin, V. V., Zwiers, F. W., Zhang, X., and Hegerl, G. C. (2007). Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate*, 20(8):1419–1444, doi:10.1175/JCLI4066.1.
- Kjellström, E., Barring, L., Jacob, D., Jones, R., Lenderink, G., and Schär, C. (2007). Modelling daily temperature extremes: recent climate and future changes over Europe. *Climatic Change*, 81(Suppl. 1):249–265, doi:10.1007/s10584-006-9220-5.
- Klein Tank, A. M. G. and Können, G. P. (2003). Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. *Journal of Climate*, 16(22):3665–3680, doi:10.1175/1520-0442(2003)016<3665:TIIODT>2.0.CO;2.
- Klein Tank, A. M. G., Zwiers, F. W., and Zhang, X. (2009). *Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation*. Climate Data and Monitoring WCDMP-No. 72, WMO-TD No. 1500. World Meteorological Organization, Geneva.
- Klok, E. J. and Klein Tank, A. M. G. (2009). Updated and extended European dataset

- of daily climate observations. *International Journal of Climatology*, 29(8):1182–1191, doi:10.1002/joc.1779.
- Knowlton, K., Lynn, B., Goldberg, R. A., Rosenzweig, C., Hogrefe, C., Rosenthal, J. K., and Kinney, P. L. (2007). Projecting Heat-Related Mortality Impacts Under a Changing Climate in the New York City Region. *American Journal of Public Health*, 97(11):2028–2034, doi:10.2105/AJPH.2006.102947.
- Koppe, C., Kovats, S., Jendritzky, G., and Menne, B. (2004). *Heat-waves: risks and responses*. Number 2 in Health and Global Environmental Change Series. World Health Organization.
- Kosatsky, T. (2005). The 2003 European heat waves. *Eurosurveillance*, 10:Available at <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=552>.
- Kovats, R. and Hajat, S. (2008). Heat stress and public health: a critical review. *Annual Review of Public Health*, 29:41–55, doi:10.1146/annurev.publhealth.29.020907.090843.
- Kovats, R. S., Hajat, S., and Wilkinson, P. (2004). Contrasting patterns of mortality and hospital admissions during hot weather and heat waves in Greater London, UK. *Occupational And Environmental Medicine*, 61(11):893–898, doi:10.1136/oem.2003.012047.
- Kovats, R. S. and Jendritzky, G. (2006). Heat-waves and human health. In Menne, B. and Ebi, K. L., editors, *Climate change and adaptation strategies for human health*, chapter 4, pages 63–97. Springer.
- Kunst, A. E., Looman, C. W. N., and Mackenbach, J. P. (1993). Outdoor air-Temperature and mortality in the Netherlands - a time-series analysis. *American Journal of Epidemiology*, 137(3):331–341.
- Kyselý, J. (2008). Influence of the persistence of circulation patterns on warm and cold temperature anomalies in Europe: Analysis over the 20th century. *Global and Planetary Change*, 62(1-2):147–163, doi:10.1016/j.gloplacha.2008.01.003.

- Laaidi, M., Laaidi, K., and Besancenot, J.-P. (2006). Temperature-related mortality in France, a comparison between regions with different climates from the perspective of global warming. *International Journal of Biometeorology*, 51(2):145–153, doi:10.1007/s00484-006-0045-8.
- Laschewski, G. and Jendritzky, G. (2002). Effects of the thermal environment on human health: An investigation of 30 years of daily mortality data from SW Germany. *Climate Research*, 21(1):91–103, doi:10.3354/cr021091.
- Lee, D. H. K. (1980). Seventy-five years of searching for a heat index. *Environmental Research*, 22(2):331–356, doi:10.1016/0013-9351(80)90146-2.
- Leith, N. A. (2007). *Single-site rainfall generation under scenarios of climate change*. PhD thesis, Department of Statistical Science, University College London.
- Leung, L. R., Mearns, L. O., Giorgi, F., and Wilby, R. L. (2003). Regional climate research: Needs and opportunities. *Bulletin of the American Meteorological Society*, 84(1):89–95, doi:10.1175/BAMS-84-1-89.
- McGregor, G. R., Pelling, M., Wolf, T., and Gosling, S. (2007). *The social impacts of heat waves*. Science Report - SC20061/SR6, Environmental Agency, UK.
- McGuffie, K. and Henderson-Sellers, A. (2005). *A Climate Modelling Primer*. John Wiley & Sons, 3rd edition.
- 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., and Nikiforov, B. (2008). International study of temperature, heat and urban mortality: the 'ISOTHURM' project. *International Journal of Epidemiology*, 37(5):1121–1131, doi:10.1093/ije/dyn086.
- McMichael, A. J., Woodruff, R. E., and Hales, S. (2006). Climate change and human health: present and future risks. *Lancet*, 367(9513):859–869, doi:10.1016/S0140-6736(06)68079-3.

- Mearns, L. O., Giorgi, F., Whetton, P., Pabon, D., Hulme, M., and Lal, M. (2003). *Guidelines for Use of Climate Scenarios Developed from Regional Climate Model Experiments*. Guidance notes from IPCC Task Group on Scenarios for Climate Impact Assessment.
- Meehl, G. A., Stocker, T. F., Collins, W. D., Friedlingstein, P., Gaye, A. T., Gregory, J. M., Kitoh, A., Knutti, R., Murphy, J. M., Noda, A., Raper, S. C. B., Watterson, I. G., Weaver, A. J., and Zhao, Z. (2007). Global Climate Projections. In Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H. L., editors, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 10, pages 747–845. Cambridge University Press.
- Michelon, T., Magne, P., and Simon-Delavelle, F. (2005). Lessons of the 2003 heat-wave in France and action taken to limit the effect of future heat-waves. In Kirch, W., Menne, B., and Bertollini, R., editors, *Extreme weather events and public health response*, chapter 13, pages 131–140. Springer.
- Michelozzi, P., Kirchmayer, U., Katsouyanni, K., Biggeri, A., McGregor, G., Menne, B., Kassomenos, P., Anderson, H. R., Baccini, M., Accetta, G., Analytis, A., and Kosatsky, T. (2007). Assessment and prevention of acute health effects of weather conditions in Europe, the PHEWE project: background, objectives, design. *Environmental Health*, 6:12, doi:10.1186/1476-069X-6-12.
- Miller, N. L., Hayhoe, K., Jin, J., and Auffhammer, M. (2008). Climate, extreme heat, and electricity demand in California. *Journal of Applied Meteorology and Climatology*, 47(6):1834–1844, doi:10.1175/2007JAMC1480.1.
- Moberg, A. and Jones, P. D. (2004). Regional climate model simulations of daily maximum and minimum near-surface temperatures across Europe compared with observed station data 1961-1990. *Climate Dynamics*, 23(7-8):695–715, doi:10.1007/s00382-004-0464-3.

- Moberg, A. and Jones, P. D. (2005). Trends in indices for extremes in daily temperature and precipitation in central and western Europe, 1901-99. *International Journal of Climatology*, 25(9):1149–1171, doi:10.1002/joc.1163.
- Murphy, J. M. (1999). An evaluation of statistical and dynamical techniques for downscaling local climate. *Journal of Climate*, 12(8):2256–2284, doi:10.1175/1520-0442(1999)012<2256:AEOSAD>2.0.CO;2.
- Murphy, J. M., Booth, B. B. B., Collins, M., Harris, G. R., Sexton, D. M. H., and Webb, M. J. (2007). A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philosophical Transactions of the Royal Society A - Mathematical, Physical & Engineering Sciences*, 365(1857):1993–2028, doi:10.1098/rsta.2007.2077.
- Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., and Collins, M. (2004). Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, 430:768–772, doi:10.1038/nature02771.
- Murphy, J. M., Sexton, D. M. H., Jenkins, G. J., Boorman, P. M., Booth, B. B. B., Brown, C. C., Clark, R. T., Collins, M., Harris, G. R., Kendon, E. J., Betts, R. A., Brown, S. J., Howard, T. P., Humphrey, K. A., McCarthy, M. P., McDonald, R. E., Stephens, A., Wallace, C., Warren, R., Wilby, R., and Wood, R. A. (2009). *UK Climate Projections Science Report: Climate change projections*. Met Office Hadley Centre, Exeter, UK.
- Nakićenović, N. and Swart, R. (2000). *Emission scenarios. A special report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press, New York.
- Nelder, J. A. and Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society Series A (General)*, 135(3):370–384.
- New, M., Lopez, A., Dessai, S., and Wilby, R. (2007). Challenges in using probabilistic climate change information for impact assessments: an example from the water sector.

- Philosophical Transactions of the Royal Society A - Mathematical, Physical & Engineering Sciences*, 365(1857):2117–2131, doi:10.1098/rsta.2007.2080.
- Nogueira, P., Falcão, J. M., Contreiras, M. T., Paixão, E., Brandão, J., and Batista, I. (2005). Mortality in Portugal associated with the heat wave of August 2003: Early estimation of effect, using a rapid method. *Eurosurveillance*, 10:Available at <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=553>.
- Nogueira, P. and Paixão, E. (2008). Models for mortality associated with heatwaves: update of the Portuguese heat health warning system. *International Journal of Climatology*, 28(4):545–562, doi:10.1002/joc.1546.
- Noguer, M., Jones, R., and Murphy, J. (1998). Sources of systematic errors in the climatology of a regional climate model over Europe. *Climate Dynamics*, 14(10):691–712, doi:10.1007/s003820050249.
- Oke, T. R. (1987). *Boundary Layer Climates*. Routledge, 2nd edition.
- O’Neill, B. C. (2004). Conditional probabilistic population projections: an application to climate change. *International Statistical Review*, 72(2):167–184, doi:10.1111/j.1751-5823.2004.tb00231.x.
- O’Neill, M. S., Zanobetti, A., and Schwartz, J. (2003). Modifiers of the temperature and mortality association in seven US cities. *American Journal of Epidemiology*, 157(12):1074–1082, doi:10.1093/aje/kwg096.
- Parsons, K. (2003). *Human thermal environments: The effects of hot, moderate and cold environments on human health, comfort and performance*. CRC Press, 2nd edition.
- Pattenden, S., Nikiforov, B., and Armstrong, B. G. (2003). Mortality and temperature in Sofia and London. *International Journal of Biometeorology*, 57(8):628–633, doi:10.1136/jech.57.8.628.
- Pauli, F. and Rizzi, L. (2008). Summer temperature effects on deaths and hospital admissions among the elderly population in two Italian cities. *Journal of Applied Statistics*, 35(3):263–276, doi:10.1080/02664760701833354.

- Peng, R. D., Dominici, F., and Louis, T. A. (2006). Model choice in time series studies of air pollution and mortality. *Journal of the Royal Statistical Society Series A - Statistics in Society*, 169(Part 2):179–198, doi:10.1111/j.1467-985X.2006.00410.x.
- Pierce, D. A. and Schafer, D. W. (1986). Residuals in Generalized Linear Models. *Journal of the American Statistical Association*, 81(396):977–986.
- Pirard, P., Vandentorren, S., Pascal, M., Laaidi, K., Le Tertre, A., Cassadou, S., and Ledrans, M. (2005). Summary of the mortality impact assessment of the 2003 heat wave in France. *Eurosurveillance*, 10:Available at <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=554>.
- Pope, V. D., Gallani, M. L., Rowntree, P. R., and Stratton, R. A. (2006). The impact of new physical parametrization in the Hadley Centre climate model: HadAM3. *Climate Dynamics*, 16(2-3):123–146, doi:10.1007/s003820050009.
- R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Räisänen, J. and Ruokolainen, L. (2008). Estimating present climate in a warming world: a model-based approach. *Climate Dynamics*, 31(5):573–585, doi:10.1007/s00382-007-0361-7.
- Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A., and Taylor, K. (2007). Climate Models and Their Evaluation. In Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., editors, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 8, pages 589–662. Cambridge University Press.
- Roberts, S. (2004). Interactions between particulate air pollution and temperature in air pollution mortality time series studies. *Environmental Research*, 96(3):328–337, doi:10.1016/j.envres.2004.01.015.

- Robinson, P. (2001). On the definition of a heat wave. *Journal of Applied Meteorology*, 40(4):762–775, doi:10.1175/1520-0450(2001)040<0762:OTDOAH>2.0.CO;2.
- Rougier, J. (2007). Probabilistic inference for future climate using an ensemble of climate model evaluations. *Climatic Change*, 81(3-4):247–264, doi:10.1007/s10584-006-9156-9.
- Rowell, D. P. (2006). A demonstration of the uncertainty in projections of UK climate change resulting from regional model formulation. *Climatic Change*, 79(3-4):243–257, doi:10.1007/s10584-006-9100-z.
- Rowell, D. P. and Jones, R. G. (2006). Causes and uncertainty of future summer drying over Europe. *Climate Dynamics*, 27(2-3):281–299, doi:10.1007/s00382-006-0125-9.
- Saez, M., Sunyer, J., Castellsagué, J., Murillo, C., and Antó, J. M. (1995). Relationship between weather temperature and mortality - A time-series analysis approach in Barcelona. *International Journal of Epidemiology*, 24(3):576–582.
- Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A., and Appenzeller, C. (2004). The role of increasing temperature variability in European summer heatwaves. *Nature*, 27:332–336, doi:10.1038/nature02300.
- Schwartz, J. (2000). The distributed lag between air pollution and daily deaths. *Epidemiology*, 11(3):320–326.
- Sheridan, S. C. (2002). The redevelopment of a weather-type classification scheme for North America. *International Journal of Climatology*, 22(1):51–68, doi:10.1002/joc.709.
- Sheridan, S. C. and Kalkstein, L. S. (2004). Progress in heat watch-warning system technology. *Bulletin of the American Meteorological Society*, 85(12):1931–1941, doi:10.1175/BAMS-85-12-1931.
- Smoyer-Tomic, K. E. and Rainham, D. G. C. (2001). Beating the heat: Development and evaluation of a Canadian hot weather health-response plan. *Environmental Health Perspectives*, 109(12):1241–1248.

- Stainforth, D. A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D. J., Kettleborough, J. A., Knight, S., Martin, A., Murphy, J. M., Piani, C., Sexton, D., Smith, L. A., Spicer, R. A., Thorpe, A. J., and Allen, M. R. (2005). Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, 433(7024):403–406, doi:10.1038/nature03301.
- Steadman, R. G. (1979a). Assessment of sultriness 1. Temperature-humidity index based on human physiology And clothing science. *Journal of Applied Meteorology*, 18(7):861–873, doi:10.1175/1520-0450(1979)018<0861:TAOSPI>2.0.CO;2.
- Steadman, R. G. (1979b). Assessment of sultriness 2. Effects of wind, extra radiation and barometric pressure on apparent temperature. *Journal of Applied Meteorology*, 18(7):874–885, doi:10.1175/1520-0450(1979)018<0874:TAOSPI>2.0.CO;2.
- Steadman, R. G. (1984). A universal scale of apparent temperature. *Journal of Climate and Applied Meteorology*, 23(12):1674–1687, doi:10.1175/1520-0450(1984)023<1674:AUSOAT>2.0.CO;2.
- Stedman, J. R. (2004). The predicted number of air pollution related deaths in the UK during the August 2003 heatwave. *Atmospheric Environment*, 38(8):1087–1090, doi:10.1016/j.atmosenv.2003.11.011.
- Stephenson, D. B. (2008). Definiton, diagnosis, and origin of extreme weather and climate events. In Diaz, H. and Murnane, R., editors, *Climate extremes and society*, chapter 1, pages 11–23. Cambridge University Press.
- Stott, P. A., Gillett, N. P., Hegerl, G. C., Karoly, D. J., Stone, D. A., Zhang, X., and Zwiers, F. (2010). Detection and attribution of climate change: a regional perspective. *Wiley Interdisciplinary Reviews: Climate Change*, 1(2):192–211, doi:10.1002/wcc.34.
- Tan, J. G., Zheng, Y. F., Song, G. X., Kalkstein, L. S., Kalkstein, A. J., and Tang, X. (2007). Heat wave impacts on mortality in Shanghai, 1998 and 2003. *International Journal of Biometeorology*, 51(3):193–200, doi:10.1007/s00484-006-0058-3.

- Tebaldi, C. and Knutti, R. (2007). The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A - Mathematical, Physical & Engineering Sciences*, 365(1857):2053–2075, doi:10.1098/rsta.2007.2076.
- Tebaldi, C., Smith, R. L., Nychka, D., and Mearns, L. O. (2005). Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multimodel ensembles. *Journal of Climate*, 18(10):1524–1540, doi:10.1175/JCLI3363.1.
- Trigo, R. M., Garcia-Herrera, R., Díaz, J., Trigo, I. F., and Valente, M. A. (2005). How exceptional was the early August 2003 heatwave in France? *Geophysical Research Letters*, 32(10):L10701, doi:10.1029/2005GL022410.
- UK Met Office (2006). MIDAS Land Surface Stations data (1853-current). Available from British Atmospheric Data Centre at <http://badc.nerc.ac.uk/data/ukmo-midas>. Accessed 1 April 2010.
- UK Office of National Statistics (2006). Estimated daily mortality during July 2006 in England and Wales. *Health Statistics Quarterly*, 32 (Winter):107–111.
- UK Office of National Statistics (2007). *Focus on London*. Palgrave Macmillan.
- UK Office of National Statistics (2008a). Methodology guide - 2006-based SNPP. Available at <http://www.statistics.gov.uk/statbase/Product.asp?vlnk=997>. Accessed 1 April 2010.
- UK Office of National Statistics (2008b). Population estimates for UK, England and Wales, Scotland and Northern Ireland - current datasets. Available at <http://www.statistics.gov.uk/statbase/Product.asp?vlnk=15106>. Accessed 1 April 2010.
- UK Office of National Statistics (2008c). Subnational Population Projections (SNPP) for England (2006-based). Available at <http://www.statistics.gov.uk/statbase/Product.asp?vlnk=997>. Accessed 1 April 2010.
- Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen,

- S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J., Morcrette, J. J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J. (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612):2961–3012, doi:10.1256/qj.04.176.
- van der Linden, P. and Mitchell, J. F. B. (2009). *ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project*. Met Office Hadley Centre, Exeter, UK.
- Vandendorren, S. and Empeur-Bissonnet, P. (2005). Health impacts of the 2003 heat-wave in France. In Kirch, W., Menne, B., and Bertolini, R., editors, *Extreme weather events and public health response*, chapter 8, pages 81–87. Springer.
- Vaneckova, P., Beggs, P. J., de Dear, R. J., and McCracken, K. W. J. (2008). Effect of temperature on mortality during the six warmer months in Sydney, Australia, between 1993 and 2004. *Environmental Research*, 108(3):361–369, doi:10.1016/j.envres.2008.07.015.
- Vidal, J.-P. and Wade, S. (2008). A framework for developing high-resolution multi-model climate projections: 21st century scenarios for the UK. *International Journal of Climatology*, 28(7):843–858, doi:10.1002/joc.1593.
- Wahba, G. (1990). *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics.
- Wetherald, R. T. and Manabe, S. (1995). The mechanisms of summer dryness induced by global warming. *Journal of Climate*, 8(12):3096–3108, doi:10.1175/1520-0442(1995)008<3096:TMOSDI>2.0.CO;2.
- Wetherald, R. T. and Manabe, S. (1999). Detectability of summer dryness caused by greenhouse warming. *Climatic Change*, 43(3):495–511, doi:10.1023/A:1005499220385.

- Wetherald, R. T. and Manabe, S. (2002). Simulation of hydrologic changes associated with global warming. *Journal of Geophysical Research*, 107(D19):4379, doi:10.1029/2001JD001195.
- Wilby, R. L. (2003). Past and projected trends in London's urban heat island. *Weather*, 58(7):251–260, doi:10.1256/wea.183.02.
- Wilby, R. L. (2007). A review of climate change impacts on the built environment. *Built Environment*, 33(1):31–45, doi:10.2148/benv.33.1.31.
- Wilby, R. L., Charles, S. P., Zorita, E., Timbal, B., Whetton, P., and Mearns, L. O. (2004). *Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods*. Guidance notes from IPCC Task Group on Scenarios for Climate Impact Assessment.
- Wilby, R. L. and Wigley, T. M. L. (1997). Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography*, 21(4):530–548, doi:10.1177/030913339702100403.
- Wilks, D. S. (2006). *Statistical methods in the atmospheric sciences*, volume 91 of *International Geophysical Series*. Academic Press, 2nd edition.
- Wilks, D. S. and Wilby, R. L. (1999). The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, 23(3):329–357, doi:10.1177/030913339902300302.
- Wood, S. N. (2006). *Generalized additive models: An introduction with R*. Chapman & Hall & CRC.
- World Health Organization (2008). *Heat-health Action Plans: Guidance*. World Health Organization Regional Office for Europe.
- World Meteorological Organization (1989). *Calculation of monthly and annual 30-year standard normals*. Climate Data and Monitoring WCDMP-No. 10, WMO-TD No. 341. Geneva.

Zanobetti, A., Wand, M. P., Schwartz, J., and Ryan, L. M. (2000). Generalized additive distributed lag models: quantifying mortality displacement. *Biostatistics*, 1(3):279–292.