
Modelling extreme temperatures in Ireland under global warming using a hybrid peak-over-threshold and generalised Pareto distribution approach

Yassin Z. Osman*

MWH UK Ltd,
Dominion House,
Temple Court, Birchwood,
Warrington, WA3 6GD, UK
Email: yassin.osman@mwhglobal.com
Email: hagyassin@yahoo.co.uk
*Corresponding author

Rowan Fealy and John C. Sweeney

Department of Geography,
National University of Ireland Maynooth,
Maynooth, Co. Kildare, Ireland
Email: rowan.fealy@nuim.ie
Email: John.sweeney@nuim.ie

Abstract: In the present paper, modelling extreme temperature to assess risk of global warming in Ireland is addressed. The approach used is a combination of peak-over-threshold (POT) – generalised Pareto distribution (GPD) in which the parameter of distribution is allowed to vary with a dominant feature of climate at the location. The dominant climatic feature at a location is approximated by climatic variables derived from the National Centre for Environmental Prediction (NCEP) reanalysis data. Data from six stations were used to develop seasonal models for winter, spring, summer and autumn. Future changes in extreme temperature values are generated by using climate variables derived from (HadCM3) GCM for the A2 emissions scenario. The software extRemes was used to develop the models as it serves the proposed modelling approach. Results indicate that significant changes in extreme temperature events are projected to occur in Ireland over the course of the present century. These include hotter summers and mild winters.

Keywords: extreme temperature; generalised Pareto distribution; GPD; climate change; peak-over-threshold; POT; Ireland.

Reference to this paper should be made as follows: Osman, Y.Z., Fealy, R. and Sweeney, J.C. (2015) 'Modelling extreme temperatures in Ireland under global warming using a hybrid peak-over-threshold and generalised Pareto distribution approach', *Int. J. Global Warming*, Vol. 7, No. 1, pp.21–47.

Biographical notes: Yassin Z. Osman is a senior member of the MWH UK Ltd modelling team in Warrington for the last eight years where he provides modelling and engineering services in the areas of integrated catchment water quality modelling, wastewater networks and flood risk management. Prior to that, he had career in academia for 15 years which involved: lecturing on

engineering hydrology, civil engineering hydraulics and environmental engineering in the Republic of Ireland and Sudan; and research activities involving hydrological modelling, hydro-climatological modelling and El Nino and its impacts on East Africa hydrological systems. He has published widely on aspects of hydrological modelling and climate change with particular reference to Sudan and Ireland. His particular expertise is in hydrological modelling, climate change impacts and adaptation modelling and sustainable urban drainage systems.

Rowan Fealy is a Research-Lecturer in the Department of Geography, National University of Ireland, Maynooth where he lectures on both theoretical and applied aspects of the climate system. His particular area of expertise is in regional climate modelling and the uncertainties associated with 'downscaling' global climate scenarios to the regional or local level. Over the past 12 years, he has published widely on all aspects of climate and climate change, with particular emphasis on Ireland. In 2012, he was awarded a Fulbright Scholar in Science and Technology 2012–2013 which he undertook as a visiting scientist at the National Centre for Atmospheric Research (NCAR), Boulder, Colorado.

John C. Sweeney is Director of the Irish Climate Analysis and Research UnitS (ICARUS) and Professor of Geography at the National University of Ireland Maynooth. He has over 35 years experience in researching climate change impacts and in coordinating research projects in these areas. Currently, he is working on the use of regional climate models for analysis of current and potential future forest pest impacts. He fulfils representative functions on many national and international bodies and is also current President of the National Trust for Ireland.

1 Introduction

The occurrence of extreme, whether associated with intense precipitation, high or low temperature extremes or storm activity, can have significant impacts on society. With increasing anthropogenic influence evident on the climate system, such events are projected by the IPCC (2007) to increase over the coming century. In order to assess and manage any associated risk resulting from the occurrence of such extreme events – use of risk management models, which include associated loss events, are required (Wu and Olson, 2009). An assessment of future extreme events (e.g., extreme precipitation or temperature), needs to be tackled within a model, which could be analytical, advanced or a complex statistical one (Wu and Olson, 2010); and it will form part of the risk management model. Therefore, adequate modelling of these extreme events, and associated return periods, is of great importance, and ultimately lead to a better assessment and management of the associated risks.

Projection of future climate variables are traditionally obtained from general circulation models (GCMs) of the atmosphere, which usually provides output at a relatively coarse resolution (typical grid scale is 300 km × 300 km); whereas flood or heat wave impact studies are typically required to be taken at sub-grid scale (of order 10–50 km²). Thus, there is a disjoint between the scale at which GCMs output future climate information and the scale at which the events of interest occur. Therefore, a methodological step is required to link the large scale GCM variables to a finer scale at

which temperature or rainfall are normally measured. The model proposed here is in a form of a statistical distribution whose parameters are a function of the large scale climate; the scale at which models typically generate output. Therefore, the motivation for this research is to develop a model, under a stochastic framework, which is capable of modelling extreme temperature events for maximum (Tmax) and minimum (Tmin) temperatures, on the basis of the large scale climate.

This section introduces the context under which the present research is undertaken. Section 2 reviews techniques and methods used in modelling extreme events generally and with particular reference to Ireland. Section 3 describes the approach used in modelling extremal distributions and the software used. Section 4 describes data used in the study. Section 5 explains how the study is conducted and the steps involved in developing the temperature models. In Section 6, obtained results are presented and discussed. Summary and concluding remarks of the study are given in Section 7.

2 Review of models

Numerous techniques have been previously employed (e.g., Wilby and Dawson, 2007) to link GCM outputs to local and regional scale climate variables such as temperature and precipitation. These include dynamic regional modelling, pattern scaling, delta change methods and statistical downscaling models. The latter has become widely used in generating climate scenarios of extreme events, due to its ease of implementation and limited computational requirements. Statistical models are based on developing statistical relationships between observed local variables and the large scale state of the atmosphere. The derived relationships are then applied to a similar suite of variables produced by a GCM to generate future scenarios of the local climate variable (Karl et al., 1990; von Storch and Zwiers, 1999). In developing these statistical relationships (or transfer functions), both linear (e.g., multiple regression) and non-linear (e.g., neural networks; generalised linear models) approaches are widely used (Wilby et al., 1998). The fundamental assumptions in the statistical models are that these statistical relations are time invariant and their parameters are stationary. While such an assumption cannot be fully verified, Charles et al. (1999) suggests that the assumption of time invariance in predictor-predict and relations may be robust provided that the choice of predictors is sensible.

Departure from the fundamental assumptions of time invariant and stationary parameters in downscaling has also been attempted. For example, consideration for the non-stationary nature of surface weather variables, such as temperature and rainfall has become more prevalent, especially after a provocative statement by Milly et al. (2008). Milly et al. (2008) claim that anthropogenically-induced climate change is the reason that stationarity has died and 'cannot be revived'. Although they acknowledge that the validity of the assumption has been questioned regularly in the past, Milly et al. (2008) highlight a pressing need to address this issue due to a convergence of observations and research findings that demonstrates the urgency of the influence of climate change and variability on surface climate variables processes.

Khaliq et al. (2011) applied a non-stationary form of Poisson process to model hot weather events (HWE) in Canada in order to develop useful probabilistic climate change information. The Poisson process was applied to temperature exceeding a defined

threshold in order to differentiate between maximum and minimum daily temperature. Khaliq et al. (2011) employed a compound Poisson process gamma (CPPG) model consisting of two basic components: one relating to the occurrence of HWEs and the second, to their durations. Utility of this approach was demonstrated by using two sets of HWE data from Southern Quebec for the period 1941–2000.

Northrop and Jonathan (2011) presented a method to model non-stationary extreme values by allowing a change in a threshold used in extracting the extreme values themselves. The threshold is allowed to vary linearly with a suitable covariate which has significant impact on quantiles calculated from the extreme value series. The method was successfully illustrated using storm peak wave heights selected from 72 sites in the Gulf of Mexico.

Frías et al. (2012) analysed changes of maximum temperatures in Europe using two regional climate models (RCMs) from the EU ENSEMBLES project. Extremes were expressed in terms of return values using a time-dependent generalised extreme value (GEV) model fitted to monthly maxima. The study used data in period (1961–2000) as calibration/validation period, and assessed the changes projected for the period 2061–2100 considering the A1B emission scenario. The maximum temperature response to increased greenhouse gases, as projected by the A1B scenario, was found consistent with the two RCMs used.

The dependence of extremal model parameters on covariates has previously been considered in Ireland in the context of modelling rainfall occurrence, but not for temperature. For example, Khaliq and Cunnane (1996) modelled point rainfall occurrences with a modified Bartlett-Lewis rectangular model. They applied a six-parameter version of the model to long hourly rainfall data recorded at Valentia and Shannon Airport in Ireland. The authors applied five different sets of statistics of rainfall data for each month to estimate six parameters of the model, employing the Rosenbrock (1960) optimisation technique. Stability and sensitivity for the obtained parameters to number and type of rainfall statistics in a set were examined and an optimum set and number was derived. No use was made for climate variables in this model. The conditional distributions of rainfall depth obtained from the model compared favourably with the historical ones. Another study, undertaken by Demissie (2004) in a study of the effects of climate change on rainfall characteristics, employed atmospheric circulation and moisture variables from both the NCEP/NCAR reanalysis data and the HadCM3 GCM to model rainfall properties at Shannon, Mullingar and Rosslare synoptic stations. Initially, a statistical downscaling model was developed using both multiple linear regression and neural network models to predict local mean rainfall at these stations. Using the cluster point process model, Demissie (2004) then developed a stochastic model for simulating future extreme rainfall events in these stations by conditioning the parameters of this conceptual rainfall model upon the statistically downscaled mean rainfall properties obtained earlier. Results from the model suggested an increase in rainfall magnitude and in dry spell durations and a decrease in frequencies of rainfall depth. Kiely (1999) also investigated the impacts of climate change on precipitation and stream flow. He analysed five decades of hourly precipitation (at eight sites) and daily streamflow at four rivers in Ireland. In part of his study, he associated the trend changes in rainfall and streamflow with changes in the North Atlantic Oscillation (NAO) index that occurred in the mid-1970s.

In all the above-mentioned attempts of modelling extreme events in Ireland, only rainfall was considered and no study included extreme temperature events. Although the probabilistic nature and seasonality of extreme rainfall have been acknowledged in those studies, none of the previous studies have explicitly conditioned or associated change in the parameters of extreme rainfall on a climate variable (or a covariate) or over time. Therefore, the present study seeks to fill this gap.

Previous studies of modelling extreme rainfall or temperature in which model parameters are allowed to change with time or climate variables (covariates), are found in work of Katz (1999), Coles (2001) and Katz et al. (2002). Katz et al. (2002), based on earlier work by Coles (2001), presented a methodology for statistical downscaling of extreme events through the incorporation of covariates into the extremal distribution. The developed methodology fits extremal distributions by maximum likelihood (ML), similar to the situation with time dependent parameters, but unlike a deterministic trend variable, a covariate is itself a random variable. Therefore, by fitting the extremal distribution conditional on the values assumed by the covariate, the problem reduces to that of a time varying parameter. For instance, given the value of a covariate (y), the conditional distribution of the extremal series could be assumed to follow a generalised extreme value (GEV) distribution with location parameter $\mu(y)$, scale parameter $\sigma(y)$ and shape parameter $\gamma(y)$. A typical parameterisation would be the same as in the following equation:

$$\begin{cases} \mu(y) = \mu_0 + \mu_1(y) \rightarrow \text{changing location with } y \\ \ln \sigma(y) = \sigma_0 + \sigma_1(y) \rightarrow \text{changing scale with } y \\ \gamma(y) = \gamma \rightarrow \text{unchanging skewness with } y \end{cases} \quad (1)$$

More generally, the covariate y could actually be a vector (i.e., consisting of one or more covariates, say y_1, y_2 , etc.).

The main factor in obtaining a good extremal model in any location depends on selection of appropriate covariate(s) that might has/have dominant effects on the local/regional scale variable on an annual or seasonal time scale. One natural candidate to serve as a covariate for hydrologic extremes is the El Nino-Southern Oscillation (ENSO) phenomenon, the dominant mode in global climate variation on an annual time scale (e.g., Katz et al., 2002). It has been associated with climate anomalies (such as droughts or floods) across large regions of the world. Similarly, the NAO, which is the dominant mode of wintertime atmospheric variability in the North Atlantic, has significant influence on climate variability in Western Europe, and specifically Ireland as highlighted by Kiely (1999).

Similar to the case of traditional deterministic downscaling, in which large-scale atmospheric variables at grid point level are the field from which input variables of the downscaling models are selected, these large-scale atmospheric variables may also have the same effects on the extremal distribution parameters. Consequently, the local/regional extremal events could be affected by a change in the pattern of the large-scale atmosphere-ocean circulation at the grid point level corresponding to it. Therefore, the large-scale atmospheric variables are considered here as local covariates which affect extremal events (e.g., extreme rainfall, extreme maximum and minimum temperature).

The methodology proposed by Coles (2001) and Katz et al. (2002) in downscaling extremal events are applied in an Irish context in the present research. However, unlike these studies, the extremal models presented here are seasonally based and their

associated covariates are selected from the large-scale atmospheric variables at a grid point level which corresponds to Ireland. The basic assumption made here is that parameters of a seasonal extremal distribution model at a location/region changes as function of large-scale atmospheric variables at the grid point level, since these variables incorporated the effects of NAO.

3 The proposed model

Generally, there are two main statistical models commonly used in modelling extreme values. These are the annual maximum, or block maxima (BM) model and peak-over-threshold (POT) model. The BM model uses a series of extreme values formed by selecting the highest value in a year or a block of time and then proceeds with fitting a statistical distribution to this extracted series. The POT model on the other hand uses all data above a threshold to form a series of extreme values and then proceeds with fitting a statistical distribution to this series. A rigorous discussion of merits and demerits of each model and the appropriate statistical distribution to be used with each one is given in Cunnane (1989), Coles (2001) and Palutikof et al. (2003) and only that part relevant to the current study is mentioned here. The description given here for the hybrid POT-GPD approach is based on the work of Gilleland et al. (2005).

The modelling concept of POT is used in the present study to model extreme values series of maximum and minimum temperature, as it contains more information than selection of the BM value. Thresholds used in extracting the POT series are determined for each site using a specific procedure which will be described in Section 3.2. The appropriate distribution normally associated with such model, as mentioned in Cunnane (1989), Coles (2001) and Palutikof et al. (2003), is any one drawn from the family of generalised Pareto distribution (GPD). The distribution function, $F(X)$, of the GPD is given by:

$$F(X) = 1 - \left[1 + \frac{\varepsilon}{\sigma}(x-u) \right]^{-1} \quad (2)$$

where x is the random variable, $x > u$; and σ is the scale parameter, $\sigma > 0$, with

u = a threshold

ε = shape parameter.

Depending on the value of the shape parameter, ε , the distribution can be classified as GPD type I, type II or exponential as follows:

- 1 if $\varepsilon > 0$, the distribution is GPD type I
- 2 if $\varepsilon < 0$, the distribution is GPD type II
- 3 if $\varepsilon = 0$, the distribution is an exponential distribution defined by:

$$F(X) = 1 - e^{-\frac{(x-u)}{\sigma}}. \quad (3)$$

The return level-return period relation, or X_T - T relation, is given by:

$$X_T = u + \frac{\sigma}{\varepsilon} [(\lambda T)^\varepsilon - 1] \quad (4)$$

for GPD type I and type II, and by:

$$X_T = u + \sigma(\lambda T) \quad (5)$$

for exponential distribution, where

$\lambda = m/n$, m is the number of peak over threshold extremes;

and n is the total number of years.

T = return period (or recurrence period) in years.

The covariate concept (Gilleland et al., 2005) is based on associating a climate variable(s), considered to affect temperature in the named location, with one or all parameters of the distribution. In the present study, similar to Katz et al. (2002), only the scale parameter is allowed to vary with dominant covariates y_1 and y_2 for investigation purposes, while the shape parameter is kept constant. This is based on the assumption that the shape parameter, a characteristic of the temperature distribution at a location, is assumed to remain constant in the current and future periods. Two functional relations for the parameter with covariates are sought here. These are:

$$\begin{aligned} \ln(\sigma(y_1, y_2)) &= \sigma_0 + \sigma_1 * y_1 + \sigma_2 * y_2 && \text{Logarithmic relation, and,} \\ \sigma(y_1, y_2) &= \sigma_0 + \sigma_1 * y_1 + \sigma_2 * y_2 && \text{Identity relation} \end{aligned} \quad (6)$$

where $\sigma(y_1, y_2) > 0$ is the new value of the scale parameter as function of the covariates, σ_0 is an intercept in the linear relation, and σ_1 and σ_2 are the slopes or trends of the variation in directions of y_1 and y_2 . In the present study, the identity relation was used to describe change in the scale parameter, since the covariates are selected using stepwise regression.

3.1 Model parameters estimation

After determining a threshold and forming the POT series, parameters of the assumed GPD are need to be estimated. One of the methods used in estimating the model parameters is the ML method. The log-likelihood function to be optimised, for $\varepsilon \neq 0$, is defined (Gilleland et al., 2005) as:

$$l(\sigma, \varepsilon) = -m \log \sigma - (1 + 1/\varepsilon) \sum_{i=1}^m \log \left(1 + \varepsilon \left(\frac{X_i - u}{\sigma} \right) \right) \quad (7)$$

when $\varepsilon = 0$ (i.e., for exponential distribution) the log-likelihood function is defined as

$$l(\sigma) = -m \log \sigma - \frac{1}{\sigma} \sum_{i=1}^m (X_i - u). \quad (8)$$

One advantage of the ML method over other methods of parameters estimation is its adaptability to changes in model structures. This allows the incorporation of model parameters when they change as function of the covariates. The above likelihood functions will, respectively, change to the following forms:

$$l(\sigma_0, \sigma_1, \sigma_2, \varepsilon) = - \sum_{i=1}^m \left\{ \log \sigma(y_{1i}, y_{2i}) - (1 + 1/\varepsilon) \log \left(1 - \varepsilon \left(\frac{X_i - u}{\sigma(y_{1i} - y_{2i})} \right) \right) \right\} \quad (9)$$

$$l(\sigma_0, \sigma_1, \sigma_2) = - \sum_{i=1}^m \left\{ \log \sigma(y_{1i}, y_{2i}) - \left(\frac{X_i - u}{\sigma(y_{1i}, y_{2i})} \right) \right\}. \quad (10)$$

As analytical maximisation of the log-likelihood function is not possible, so numerical optimisation techniques are always used for this purpose. These are generally techniques devoted for solution of non-linear equations, such as Newton-Raphson, method of scoring and BHHH method (Long, 1997). The numerical optimisation techniques of Nelder-Mead and Broyden-Fletcher-Goldfarb-Shanno (BFGS), as described in Henningsen and Toomet (2011), are employed by the ‘extRemes’ software used in this study.

3.2 Threshold selection

Selection of appropriate threshold is always difficult and represents point of weakness for a POT model over others. On one hand, a threshold must be set high enough so that only true peaks, with Poisson arrival rates (Palutikof et al., 2003) are selected. If this is not the case, the distribution of selected extremes will fail to converge to the GPD asymptote. On the other hand, the threshold must be set low enough to ensure that enough data are selected for satisfactory determination of the distribution parameters.

Accordingly, a number of procedures have been used to aid in selecting an appropriate threshold for the POT model at a site. Two of these procedures are mentioned below:

- 1 *Mean residual life graphs*: This is a plot of the mean excess over threshold as a function of threshold. For a GPD model, the graph should plot as a straight line, and the appropriate threshold value can be chosen by selecting the lowest value above which the graph is straight line (e.g., Davison, 1984).
- 2 *Model parameter graphs*: This plots estimate each parameter as a function of threshold. For a GPD model, estimates of shape parameter should be approximately constant, while estimates of scale parameter should be linear, and the appropriate threshold value can be chosen by selecting the lowest value at which the graph is straight (e.g., Coles, 2001).

In this study, the 90th percentile of the data has been used as a guide for selecting appropriate threshold for maximum temperature (Tmax) and the 10th percentile has been used as a guide for selecting appropriate threshold for minimum temperature (Tmin). Using combination of the procedures described above, thresholds guides are refined to yield appropriate ones.

3.3 Model diagnostics

As the reason for fitting a statistical model to a set of data is to draw conclusions on some aspects of the population of the observed data, such conclusions could be sensitive to the accuracy of the fitted model. Thus, it is necessary to check the model accuracy and goodness-of-fit by checking its agreement with the data that were actually used to estimate it (model descriptive ability) and also checking its ability to simulate future values (model predictive ability). Four types of model diagnostics (Gilleland et al., 2005) are used in the present study to visually check the goodness-of-fit (descriptive ability) of the GPD to model the extreme values series. These are:

- a *Probability plot*, which is a comparison of an empirical (usually percentage rank) and the fitted distribution function in equations (2) or (3). In case of perfect fit, the data would line up on the diagonal of the probability plots as will be shown in Section 5.
- b *Quantile plot*, which is also a comparison of an empirical form for estimating the exceedance and the inverse of equations (2) or (3). Any departure from linearity indicates model failure in perfectly fitting the data.
- c *Return period plot*, which shows the return period in years against the return level from equations (4) or (5). Confidence intervals can be added to the plot to increase its informativeness. Empirical estimates for the return levels are also added to the plot to be used as a model diagnostic. If the GPD model is suitable for the data, the model-based curve and empirical estimates should be in reasonable agreement.
- d *Density function plot*, which is a comparison of the probability density function of a fitted model with the histogram of the POT data. This is less informative diagnostic for model as a histogram varies substantially with the choice of grouping intervals, which makes its use difficult and subjective.

In order to assess the predictive ability of the GPD model, the method of split sample testing is used. The observed POT extreme series is divided into calibration sample (1961–1990) and validation sample (1991–2000). The distribution is fitted to the first sample using GPD without covariates and the estimated parameters are used to obtain the associated probability with which the observed POT in both samples has occurred. A second GPD fit with covariates is then performed, and the new estimated parameters are then used in conjunction with the probability obtained from the first fit to simulate model output. Correlation between the observed and simulated POT series is then established. The coefficient of determination is used here to check the model predictive ability (or efficiency) in both samples.

3.4 Choice of preferred model

When GPD parameters are considered function in covariates, there will be a number of possible models to choose from. The basic principle on choosing between models is parsimony, i.e., obtaining the simplest model (with less number of parameters) that explains as much variation in the data as possible. So in order to choose between model fits a test, known as likelihood ratio test, is used. The test proceeds as follows:

In case of two models M_0 and M_1 , where M_0 is a subset of M_1 , $M_0 \subset M_1$ (e.g., M_0 is without covariates and M_1 is with covariates), the deviance statistic is defined as:

$$D = 2\{l_1(M_1) - l_0(M_0)\} \quad (11)$$

where $l_0(M_0)$ and $l_1(M_1)$ are the maximised log-likelihoods under models M_0 and M_1 respectively. Large values of D indicate that M_1 explains substantially more of the variation in the data than M_0 ; small values of D suggest that increase in the number of model parameters does not bring worthwhile improvements to the model capacity to explain the data. Therefore, initial estimates of how large D should be before preferring model M_1 over M_0 is provided by the asymptotic distribution of the deviance function (Coles, 2001). This can be tested as follows:

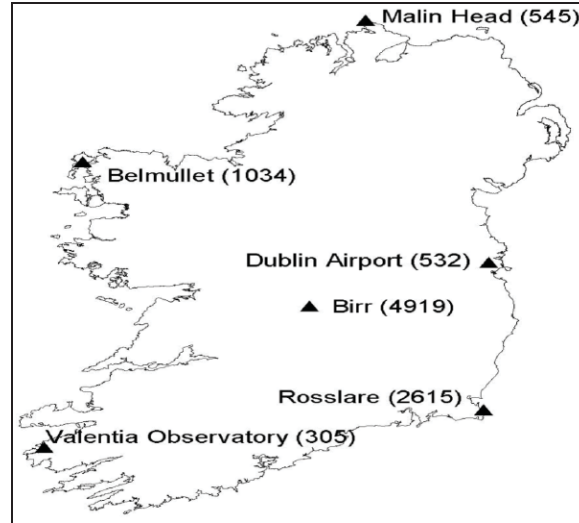
Model M_0 is rejected by a test at the α -level of significance if $D > c_\alpha$, where c_α is the $(1 - \alpha)$ quantile of the χ^2_ν distribution with ν degrees of freedom where ν is equal to the difference in the number of estimated parameters.

3.5 *The extRemes software*

The software used for fitting the GPD model to a POT series, which allows the parameter to change as function of the covariates, is extRemes version 1.62 (Gilleland et al., 2005). The software, written in the R language and benefited from Coles (2001) ‘S’ functions, is based on the concept of ML for estimating GPD parameters. The key advantage of the software is that it facilitates the fitting of statistical distributions with covariates using the ML method and has options for choosing an appropriate threshold for POT series. The Toolkit is specifically designed to facilitate the use of extreme value theory in applications oriented towards weather and climate problems that involve extremes.

4 Data

Observed daily maximum (Tmax) and minimum (Tmin) temperature data for the period 1961–2000, for six selected synoptic stations representing both coastal and inland parts of Ireland (Figure 1) are used in the present study. Data from Valentia (0305), Dublin Airport (0532), Belmullet (1034), Birr (4919), Rosslare (2615) and Malin Head (0545) synoptic stations were obtained from Met Éireann, the Irish meteorological service. Daily grid-point data for the atmospheric variables, for surface and upper atmosphere, shown in Table 1, are taken from Wilby and Dawson (2007) and they consist of National Centre for Environmental Prediction (NCEP) re-analysis data (Kalnay et al., 1996). For demonstrating the model proposed in the present study and how it can be used, climatic variables from the A2 emission scenario from the Hadley Centre’s Global Climate Model, HadCM3, extracted for the period 1961–2000, have been used. The NCEP data will serve as potential candidate of covariates for the seasonal extremal models developed.

Figure 1 Locations of synoptic stations used in the study**Table 1** Surface and atmospheric variables employed in the analysis

<i>Daily variable</i>	<i>Code</i>
Maximum temperature (K)	Tmax
Minimum temperature (K)	Tmin
Mean temperature over a day (K)	TEMP
Mean sea level pressure (hPa)	MSLP
500 hPa geopotential height	P500
800 hPa geopotential height	P800
Near surface relative humidity	RHUM
Near surface specific humidity	SHUM
Geostrophic airflow velocity	P_F
Vorticity	P_Z
Zonal velocity component	P_U
Meridional velocity component	P_V
Geostrophic airflow velocity (500 hPa)	P5_F
Vorticity (500 hPa)	P5_Z
Zonal velocity component (500 hPa)	P5_U
Meridional velocity component (500 hPa)	P5_V
Geostrophic airflow velocity (800 hPa)	P8_F
Vorticity (800 hPa)	P8_Z
Zonal velocity component (800 hPa)	P8_U
Meridional velocity component (800 hPa)	P8_V

Source: Wilby and Dawson (2004), UK SDSM data archive

5 Methodology

The steps followed to build the seasonal extreme temperature models using the hybrid POT-GPD approach are summarised in the following *seven* steps:

- 1 The daily maximum (Tmax) and minimum (Tmin) temperatures, together with the corresponding atmospheric variables obtained from NCEP and HadCM3 are arranged into four seasons. Winter is defined as December, January, and February (*DJF*); spring as March, April and May (*MAM*), summer as June, July and August (*JJA*) and autumn as September, October and November (*SON*). The lead and lag of variables in Table 1 were also derived to create additional covariate time series. A suffix of $_1$ ($_2$) is added to the present variable coding to represent a lagged time series of the variable (e.g., P_F_1, P5_U_2), whereas a suffix of +1 (+2) is added to represent a leaded time series of the variable (e.g., MSLP+1, SHUM+2).
- 2 Threshold guides for an extreme seasonal series (u) of Tmax and Tmin for each station are obtained using their 90th and 10th percentiles, respectively, as guides, to comply with extreme event indices defined in the Statistical and Regional Dynamical Downscaling of Extremes for European Regions (STARDEX Project, 2003). The threshold is then refined while fitting the model and the optimum threshold in each case, is taken when estimated values of estimated parameters stabilise. The base period for the calculation of thresholds is the period 1961–1990. This 30-year period, defined by the World Meteorological Organisation (WMO) as the 30-year normal period, is considered representative of the present day climate and encompasses a range of natural variability (IPCC, 2001). Using the threshold, seasonal series of the Tmax and Tmin extreme values are extracted for each station together with their corresponding possible set of covariates. The Tmin series is modelled here as a maximum extreme series by transforming it into a maximum series. This is done by multiplying the extreme values in the minimum extreme series by (-1) to transform the whole series into a series of maximum negative values. To maintain consistency, the threshold of the POT series should also be multiplied by (-1) when estimating the Tmin quantile. After fitting the model and estimating its parameters, the calculated quantile should finally be multiplied by (-1) to transform it back to the normal domain.
- 3 Covariate selection exercises are then run using stepwise regression between the extracted POT Tmax and Tmin series and the possible set of all covariates. Initially, a cross correlation is conducted between all possible seasonal covariates, at each station. This helps in excluding covariates demonstrating high degree of co-linearity. Then, covariates-extreme value correlations are obtained for each station using stepwise regression. This analysis, in combination with the cross correlation, allow determination of which covariates are most strongly correlated with the precipitation, which in turn helps in making an adequate selection of covariates for use in downscaling and reduces the problem of multi-co linearity. A t-test for significance of the correlation between the precipitations POT series and each covariate, obtained via the stepwise regression, is then run to help select the most dominant covariates (based on test results) in the location to use in the POT-GPD model. Tables 2 and 3 below shows, respectively, summary of Tmax and Tmin statistics, appropriate covariates, thresholds, and number of extracted extremes in each station.

Table 2 Tmax seasonal models, statistics and parameters

Seasonal model	Threshold u	m	λ	Covariates		Parameters (no cov.)			Parameters (with covariates)			Coefficient of determination (R^2)	
				y_1	y_2	σ_0	ε_0	σ_1	σ_2	ε_0	Calibration	Validation	
Tmax0305-Aut	17	420	14.00	P850	TEMP	1.677	-0.056	0.190	0.509	1.299	-0.289	0.93	0.95
Tmax0532-Aut	18	261	8.70	TEMP	P500_2	2.096	-0.323	-0.838	2.123	0.398	-0.440	0.89	0.91
Tmax0545-Aut	16	368	12.27	P500_2	SHUM_2	1.789	-0.162	0.930	0.186	0.673	-0.244	0.94	0.92
Tmax1034-Aut	17	209	6.97	P8_U	TEMP	1.500	-0.124	0.269	-0.578	0.974	-0.321	0.94	0.89
Tmax2615-Aut	17	306	10.20	TEMP	P8_F_1	1.357	-0.230	0.005	1.196	-0.116	-0.336	0.91	0.90
Tmax4949-Aut	18	254	8.47	P5_F	TEMP	2.168	-0.280	-0.695	-0.293	2.227	-0.411	0.91	0.96
Tmax0305-Spr	15	386	12.87	TEMP	RHUM_2	2.393	-0.062	1.535	1.532	-0.588	-0.312	0.91	0.91
Tmax0532-Spr	16	261	8.70	P500	TEMP_1	1.902	-0.164	1.720	0.807	1.029	-0.387	0.94	0.92
Tmax0545-Spr	14	320	10.67	RHUM	TEMP	2.296	-0.095	2.011	-0.215	1.788	-0.383	0.91	0.82
Tmax1034-Spr	15	262	8.73	RHUM	TEMP_1	2.687	-0.141	2.167	-0.455	1.886	-0.494	0.92	0.93
Tmax2615-Spr	14	331	11.03	TEMP	P850_2	1.630	-0.165	1.372	1.005	0.357	-0.341	0.91	0.92
Tmax4949-Spr	15	564	18.80	RHUM	TEMP	2.572	-0.118	1.727	-0.408	2.428	-0.377	0.92	0.95
Tmax0305-Sum	20	370	12.33	RHUM	TEMP	2.438	-0.146	-0.491	-0.425	1.498	-0.434	0.93	0.95
Tmax0532-Sum	21	365	12.17	P8_Z_2	TEMP_2	1.772	-0.161	0.408	-0.079	0.884	-0.286	0.93	0.93
Tmax0545-Sum	19	303	10.10	TEMP	P500_1	2.201	-0.230	0.796	0.589	0.350	-0.263	0.96	0.97
Tmax1034-Sum	19	335	11.17	TEMP	RHUM_1	2.371	-0.098	-0.083	1.372	-0.272	-0.334	0.95	0.91
Tmax2615-Sum	20	236	7.87	P850	TEMP	1.500	-0.136	-0.417	0.488	0.967	-0.327	0.93	0.88
Tmax4949-Sum	22	342	11.40	RHUM	TEMP	2.684	-0.227	-2.117	-0.023	2.787	-0.556	0.94	0.92
Tmax0305-Win	12	325	10.83	TEMP	P8_V_1	0.782	-0.161	0.759	0.211	0.183	-0.221	0.93	0.94
Tmax0532-Win	11	442	14.73	TEMP	P8_F_2	1.425	-0.222	1.486	1.139	0.192	-0.302	0.92	0.92
Tmax0545-Win	11	249	8.30	TEMP	P5_V_1	1.067	-0.201	1.260	0.739	0.185	-0.400	0.94	0.94
Tmax1034-Win	11	313	10.43	P8_V	TEMP	0.983	-0.322	0.904	0.159	0.193	-0.359	0.94	0.93
Tmax2615-Win	11	329	10.97	P5_Z	SHUM	0.979	-0.295	0.978	-0.068	0.365	-0.345	0.95	0.93
Tmax4949-Win	11	439	14.63	P8_U	TEMP	1.346	-0.282	1.375	0.141	0.966	-0.276	0.93	0.92

Table 3 Tmin seasonal models, statistics and parameters

Seasonal model	Threshold u	m	λ	Covariates		Parameters (no cov.)		Parameters (with covariates)				Coefficient of determination (R^2)	
				y_1	y_2	σ_0	ε_0	σ_0	σ_1	σ_2	ε_0	Calibration	Validation
Tmin0305-Aut	3	203	6.77	RHUM	TEMP_1	2.749	-0.360	1.718	-0.153	-1.641	-0.673	0.94	0.96
Tmin0532-Aut	1	198	6.60	SHUM	P_F_1	2.010	-0.435	1.775	-0.381	-0.321	-0.553	0.96	0.94
Tmin0545-Aut	3	185	6.17	RHUM	TEMP	1.822	-0.301	1.522	-0.253	-0.504	-0.433	0.94	0.94
Tmin1034-Aut	3	250	8.33	TEMP	P8_F_1	2.569	-0.449	1.806	-0.859	-0.487	-0.585	0.94	0.90
Tmin2615-Aut	3	179	5.97	P8_V_1	SHUM_1	1.794	-0.265	0.885	0.454	-1.639	-0.453	0.89	0.92
Tmin4949-Aut	0	254	8.47	P_U_1	TEMP_1	3.388	-0.430	1.667	-0.512	-1.149	-0.386	0.94	0.92
Tmin0305-Spr	1	164	5.47	P8_Z	TEMP_1	1.742	-0.207	0.665	-0.016	0.894	-0.354	0.75	0.56
Tmin0532-Spr	0	138	4.60	SHUM_1	MSLP_2	1.292	0.003	0.491	-0.748	-0.499	-0.265	0.92	0.77
Tmin0545-Spr	1	158	5.27	P8_F_1	RHUM_1	1.197	-0.108	1.069	-0.025	-0.441	-0.162	0.91	0.85
Tmin1034-Spr	1	189	6.30	P_F	TEMP	1.451	-0.093	0.114	-0.227	-1.180	-0.370	0.92	0.97
Tmin2615-Spr	2	212	7.07	P_V	TEMP	1.594	-0.260	-0.136	0.098	-1.409	-0.382	0.90	0.95
Tmin4949-Spr	0	340	11.33	RHUM	TEMP_1	1.787	-0.083	0.482	-0.244	-1.153	-0.272	0.93	0.90
Tmin0305-Sum	8	210	7.00	P_F	TEMP_1	1.618	-0.242	2.089	-0.753	-1.280	-0.485	0.91	0.95
Tmin0532-Sum	7	155	5.17	P5_V	TEMP_2	1.521	-0.206	1.867	-0.324	-0.780	-0.282	0.92	0.87
Tmin0545-Sum	8	183	6.10	P8_F	RHUM_1	1.327	-0.170	1.176	-0.121	-0.112	-0.189	0.97	0.97
Tmin1034-Sum	8	292	9.73	P8_F	TEMP_1	1.613	-0.169	2.449	-0.293	-1.079	-0.381	0.92	0.90
Tmin2615-Sum	9	277	9.23	P500_1	TEMP_1	1.482	-0.274	2.005	-0.047	-0.880	-0.380	0.93	0.96
Tmin4949-Sum	6	213	7.10	P8_U	SHUM	1.766	-0.206	1.929	-0.506	-0.813	-0.371	0.92	0.94
Tmin0305-Win	0	306	10.20	P_Z_1	TEMP_1	2.305	-0.288	0.002	-0.047	-1.402	-0.442	0.91	0.95
Tmin0532-Win	-1	311	10.37	P8_F_1	TEMP_1	1.850	-0.083	-0.040	-0.376	-1.055	-0.173	0.92	0.94
Tmin0545-Win	0	246	8.20	P8_U	TEMP_1	1.854	-0.213	-0.170	-0.223	-1.176	-0.410	0.90	0.94
Tmin1034-Win	0	370	12.33	TEMP_1	RHUM_2	2.391	-0.243	0.235	-1.188	-0.173	-0.274	0.91	0.95
Tmin2615-Win	0	188	6.27	TEMP	P5_U_1	1.493	-0.284	0.649	-0.550	-0.043	-0.412	0.94	0.97
Tmin4949-Win	-2	387	12.90	P8_F	TEMP_1	2.909	-0.174	0.299	-0.264	-1.470	-0.244	0.93	0.92

- 4 Having selected the appropriate covariates for different POT series, a GPD fitting with ML is performed in two steps to develop the seasonal extremal model in the station. In Step 1, the GPD fitting is performed using a guided threshold value and without the use of covariates. Following the second procedure described in Section 2.2 a refining value for the threshold is obtained. A refitting for the GPD is then performed again using the refined value of the threshold and the model is termed base model (M_0) for the station. In step two, a third GPD fitting is performed with use of the selected covariates for the scale parameter only, and the model is termed seasonal model (M_1) for the station. Diagnostic plots for the fit, provided by the extRemes software, judge how well model M_0 fits the data. Which of the two models (M_1 and M_0) is preferred over the other is governed by running the likelihood ratio test. Moreover, an evaluation for the coefficient of determination, R^2 , between observed and simulated extreme values series is used as a test for model predictability and performance. Extreme values series from the period 1961–1990 is used for calibration and from the period 1991–2000 is used for validation. Values of R^2 for each period and season are also shown in Tables 2 and 3.
- 5 Perturbed climate seasonal return level – return period relations for each station are then developed by extracting corresponding values of covariates from the HadCM3 model outputs. The extracted covariates are used to generate possible future values assumed by the scale parameter ($\sigma(y_1, y_2)$), using fitting parameters of model M_1 and equation (6). The shape parameter is considered constant. For each possible value of the scale parameter a value for the return level X_T for a range of return periods T years is calculated using equation (4). Return periods considered are 2, 3, 5, 7, 10, 20, 30, 50, 75 and 100 years. Values of X_T from model M_0 (referred hereinafter as NOCLM) are also calculated for the same return periods using equation (4) and model parameters values from fit M_0 .
- 6 The maximum, minimum and average values of the calculated X_T (Tmax/Tmin quantile) in each seasonal downscaling model are then obtained for the baseline period 0 (CLM1961-1990), period 1 (CLM1991-2020), period 2 (CLM2021-2050), and period 3 (CLM2051-2080). The number of years in each period (n) is 30 years.
- 7 For each X_T series in the periods outlined above, the maximum value of Tmax (or minimum value in case of Tmin) in the series is taken to represent a point in the effective X_T - T relation in that period. The max (X_T) and min (X_T) points obtained are finally plotted against return period T to yield the affective seasonal return level-return period curve of Tmax/Tmin for any of the considered periods at all stations.

6 Analysis and discussion of results

Results obtained by applying the methodology described above on data from different locations, are analysed and discussed in this section. Tabular and graphical forms of presentation are used to aid the analysis and discussion, which come here in two parts. The first part is devoted for analysing the goodness of fit of the combined POT-GPD as a model for risky events of temperature at the stations and how incorporation of covariates improves the model predictability. The second part concerns with the discussion of

how the developed seasonal models could be used to drive an effective seasonal return level – return period relation, and the usefulness of these relations as tools for risk management of, e.g., heat wave or freezing conductions resulting from global warming at these stations.

6.1 Building the Tmax/Tmin POT-GPD seasonal models

Following the steps described in Section 5, a GPD with covariates model has been employed to fit seasonal POT series of maximum (Tmax) and minimum (Tmin) temperatures at all selected stations using the extRemes software. The hybrid POT-GPD modelling approach is used to build a total of 48 different POT-GPD seasonal models for Tmax and Tmin, in which the scale parameter of the GPD is allowed to vary as a function of selected covariates (four seasonal models for Tmax and four seasonal models for Tmin at each station). Tables 2 and 3 present results of estimated model parameters for each Tmax/Tmin models. Analysis and discussion of each group of models are given below.

6.1.1 Tmax seasonal models

Temperature is considered a less problematic physical variable to model. However, extreme values of Tmax can sometimes be difficult to model. The POT series of seasonal Tmax for each station is extracted using the 90th percentile of the series as an initial threshold. The methodology described in Section 5 is followed to build the Tmax seasonal models at each station. The stepwise regression process revealed that the optimum candidates for Tmax0305-Sum model covariate in the location are RHUM and TEMP, both of which display significant correlation with the Tmax series ($r = -0.428$ for RHUM and $r = 0.455$ for TEMP). After selecting suitable threshold and appropriate covariates, the data is fitted to the GP models M_0 and M_1 . Likelihood ratio test is then used to select the preferred POT-GPD seasonal model for Tmax at the named station.

Figures 2 and 3 showed diagnostic plots of M_0 and M_1 , respectively, for the Valencia Observatory Tmax0305-Sum model. The diagnostics plots in these figures demonstrate that GPD fits the Tmax extreme values series at this station very well and that addition of covariates to the model produce a new model which also follow the same distribution. Similar results are obtained for all Tmax seasonal models in the other stations. The likelihood ratio test, described in Section 3.4, revealed that M_1 is preferred over M_0 , for this station and all other stations. Since the calculated values of D were significant at the 0.05 level, another significance level could be used.

All the Tmax POT-GPD seasonal models developed in this study are shown in Table 2. The values of shape parameter in Table 2 indicate that all the Tmax seasonal models obtained are classified as GPD II, even after inclusion of covariates. The GPD II fits extreme Tmax series very well as judged by the diagnostics plots in Figures 2 and 3. The developed seasonal models predictability and efficiency are further checked here by evaluating the value of coefficient of determination, R^2 , yielded by correlating the observed and simulated extreme series, as explained in Section 5. For a significance level of 0.05, values for the coefficient of determination are found to be very high (more than 80%) for all models for both the calibration and validation periods, as shown in the last two columns of Table 2. The reason for obtaining higher values of correlations is the inclusion of covariates in estimating future scale parameter for the distribution (i.e.,

treating the distributional model as non-stationary). Therefore, the high values of the coefficient of determination and the good agreement in the diagnostic plots suggest that these models can reasonably be used to describe change in the occurrence of extreme Tmax events at these locations.

Figure 2 Diagnostic plots for Tmax0305-Sum model at Valentia station for fit M_0 (see online version for colours)

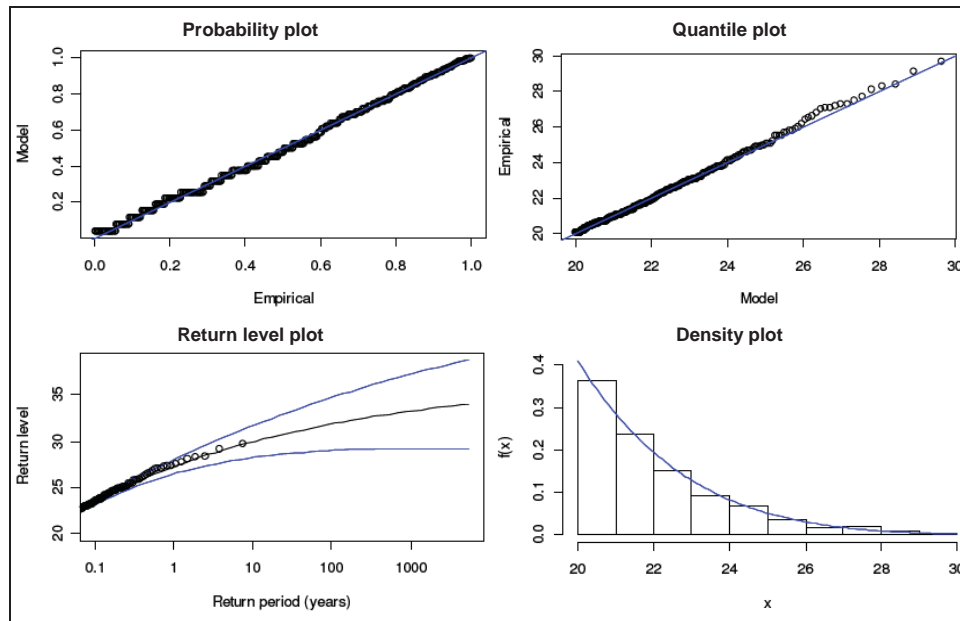
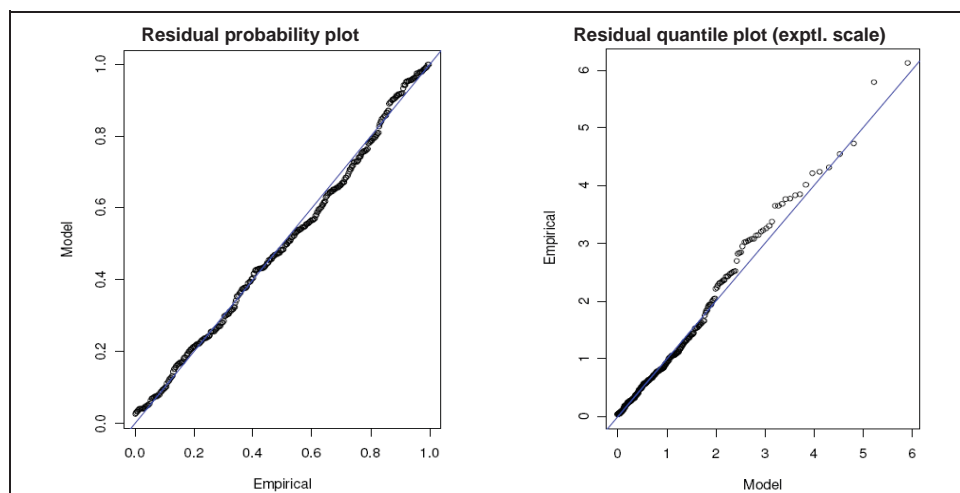


Figure 3 Diagnostic plots for Tmax0305-Sum model at Valentia station for fit M_1 (see online version for colours)



6.1.2 *Tmin seasonal models*

The extreme values series of Tmin temperature is extracted using the 10th percentile of the data series as initial threshold. The extracted series, which represent a minimum series of the variable, cannot be modelled as it is by the hybrid POT-GPD approach. The Tmin series is transformed to a maximum series following the procedure as outlined in Step 2 of Section 5. Then the rest of the methodology described in Section 5 is followed to build the Tmin seasonal models. A threshold value of 10th percentile was first used, then refined before selecting appropriate covariates from the NCEP data then go ahead with fitting models M_0 and M_1 . The likelihood ratio test is then used to choose the best extreme Tmin POT-GPD seasonal model for each station.

Figures 4 and 5 show the diagnostic plots resulted from fits M_0 and M_1 , respectively, for the Valentia Observatory Tmin0305-Win model. The diagnostic plots of Figure 4 indicate that GPD II fits the Tmin extreme values very well at this station. Similar results are also obtained for all other station models. Although the likelihood ratio tests, described in Section 3.4, revealed that M_1 is preferred over M_0 , however the diagnostic plots of Figure 5 may look a little bit confusing. The confusion here comes as these plots reflect the influences of climate variables in a quantile magnitude and frequency, which are represented here by including of these covariates in the model.

Figure 4 Diagnostic plots for Tmin0305-Win model at Valentia for fit M_0 (see online version for colours)

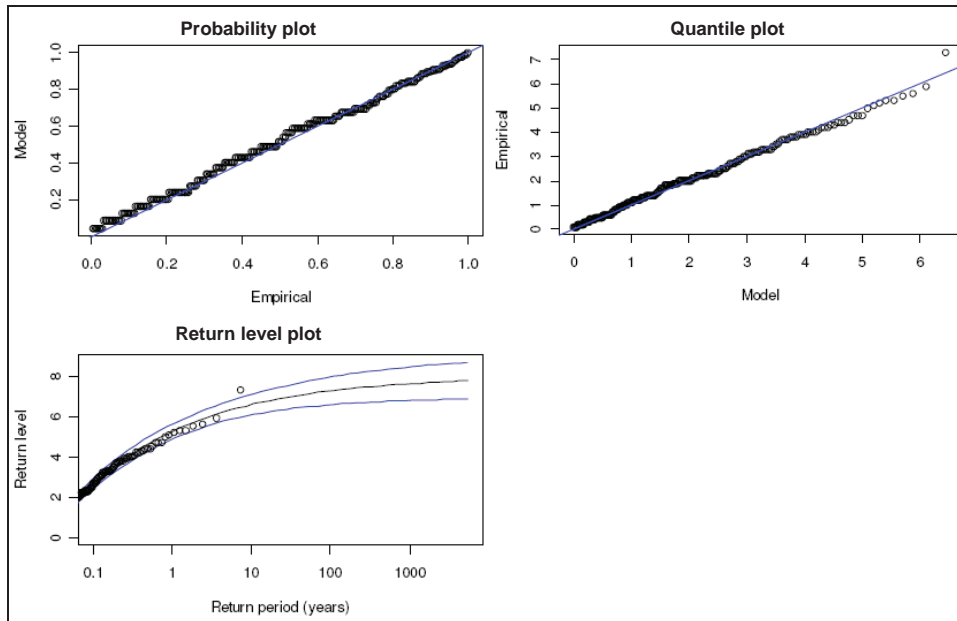


Figure 5 Diagnostic plots for Tmin0305-Win model at Valentia for fit M_1 (see online version for colours)

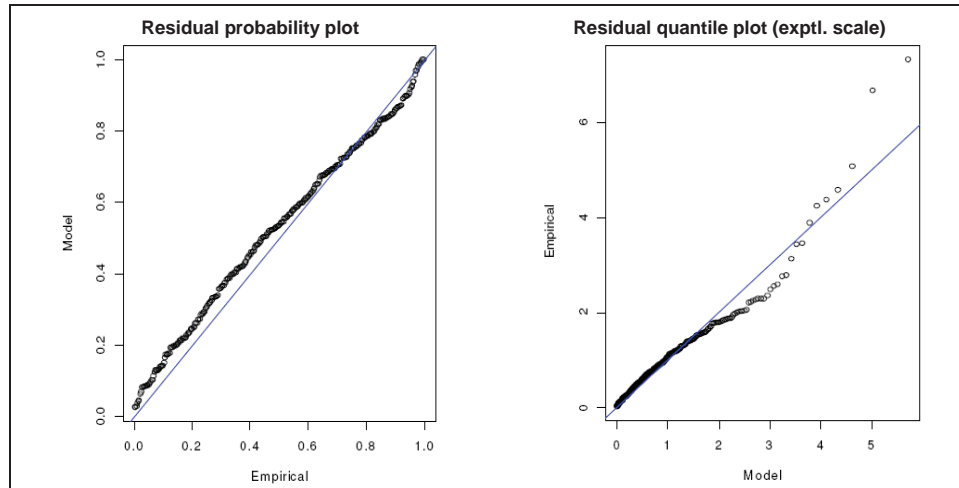


Table 3 shows all Tmin POT-GPD seasonal models developed in the present study. It is observed from these results that the shape parameter values in all models do not change their sign. That is to say, all the Tmin seasonal models follow the form of GPD II, even after inclusion of covariates in the models. It can also be noticed that GPD II predicts Tmin extreme quantiles very well as judged by the diagnostic plots in Figure 4. But the climate driven model of Tmin is not that good, judging by its fit in Figure 5. So, it can be noted here that GPD II does not predict Tmin extreme values very well. Thus, these models are not considered robust for generating future event of Tmin.

6.2 Climate driven return level – return period relations

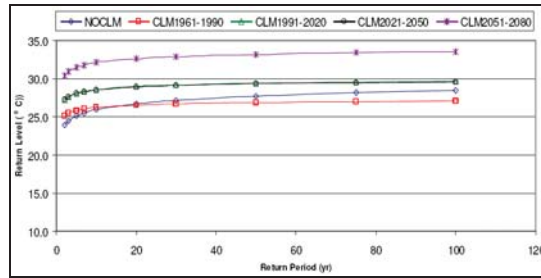
The relations analysed and discussed in this section are obtained using the seasonal models presented in Tables 2 and 3. The derived effective return level – return period relations are then used to prepare characteristic curves for each station. Samples of these curves are shown here to demonstrate how future climate may possibly affect the magnitude and frequency of extreme temperature events associated with a specified return period. The characteristic curves for Valentia and Birr stations are used here to illustrate this.

6.2.1 Tmax effective relations

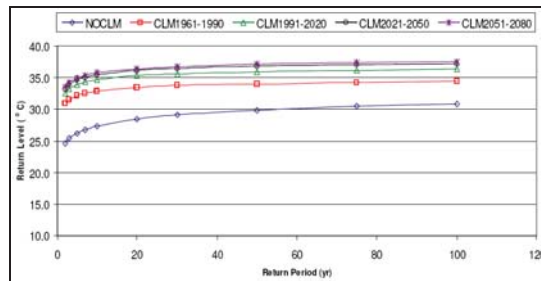
Figures 6(a), 6(b), 6(c) and 6(d) and Figures 7(a), 7(b), 7(c) and 7(d) represent graphical forms of Valentia and Birr effective Tmax return levels – return period relations. Each seasonal relation at a station shows five curves; one for each modelling period described in Section 4, and a fifth curve to represent the relation from NOCLM (or fit M_0). All relation curves at both stations show general increase in the maximum temperature quantile magnitude as time moves into future with change in climatic conditions driven by the ongoing global warming. Period 3 curves, e.g., show the highest of this increase in the quantile magnitudes while the baseline curves give the lowest for all seasons except

the summer. Similarly, the curves correspond to the NOCLM (M_0) generally fall far below the baseline curves for all seasons except the summer.

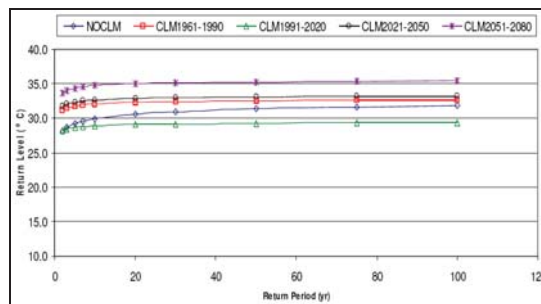
Figure 6 return level versus return period plot for (a) Tmax0305-Aut, (b) Tmax0305-Spr, (c) Tmax0305-Sum and (d) Tmax0305-Win (see online version for colours)



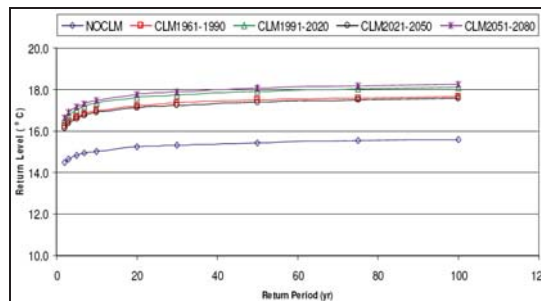
(a)



(b)

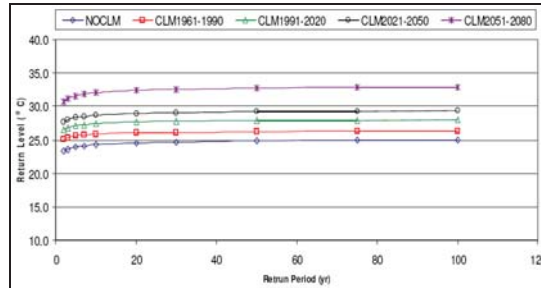


(c)

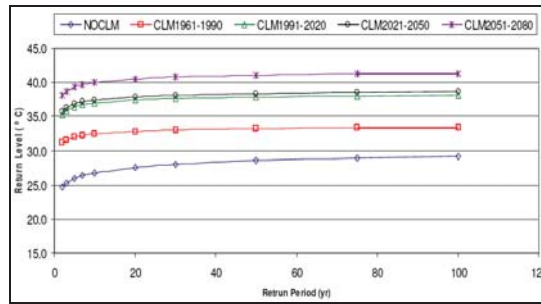


(d)

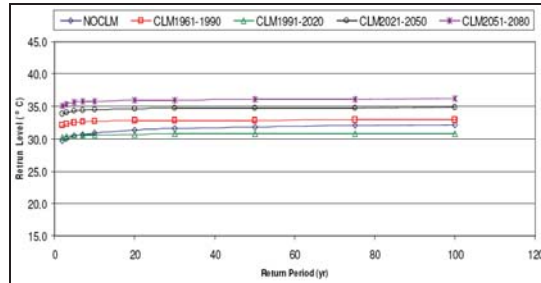
Figure 7 Return level versus return period plot for (a) Tmax4919-Aut, (b) Tmax4919-Spr, (c) Tmax4919-Sum and (d) Tmax4919-Win (see online version for colours)



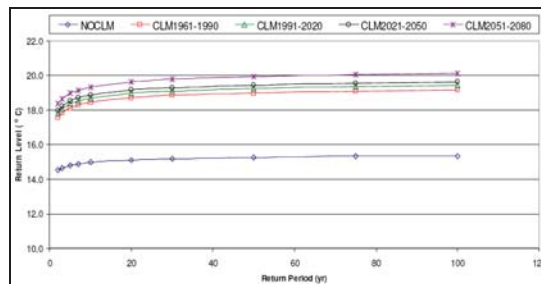
(a)



(b)



(c)



(d)

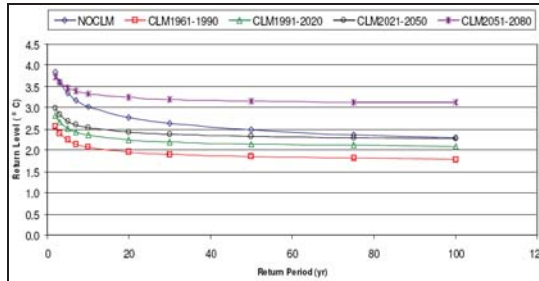
The curves in Figures 6(a), 6(b), 6(c) and 6(d) also reveal that difference in temperature increase in Tmax magnitude, as a result of climate change differs seasonally. For example, for the summer period at Valentia, a return period of 100 years gives a Tmax of 31.9°C for NOCLM (M_0) curve, 32.6°C for the baseline period, 29.3°C for period 1, 33.2°C for period 2, and 35.4°C for period 3. So, on average, a likely increase of around 2–3% is expected in the Tmax summer values as a consequence of climate change. This summer difference in temperature increase is significantly exceeded in Valentia spring and winter relations.

For Birr station effective relations shown in Figures 7(a), 7(b), 7(c) and 7(d), similar results, as those obtained for Valentia, are found. The only difference here is that climate change effects are more pronounced. All Tmax return level-return relations obtained here indicate that warmer conditions are likely for the future.

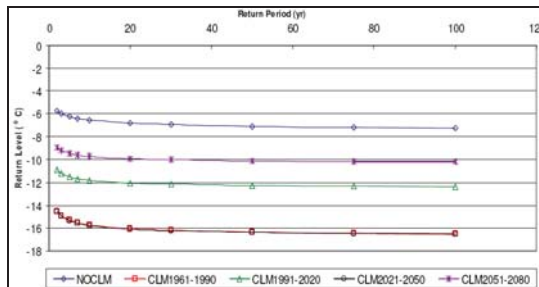
6.2.2 Tmin effective relations

Figures 8(a), 8(b), 8(c) and 8(d) and Figures 9(a), 9(b), 9(c) and 9(d) show curves of effective return level – return period relations for Tmin at Valentia and Birr stations, respectively. The curves in these figures demonstrate that a general decrease in Tmin is expected as a consequence of climate change. This would translate into an increased likelihood of severe winters at these locations. The decrease in a Tmin magnitude, of 100 years return period, is found to vary with climate periods considered. A maximum decrease of 9°C is projected at Valentia during wintertime between the baseline period and period 2. During the summer time, the decrease in Tmin is relatively lower. Consequently, colder conditions are expected at this location.

Figure 8 Return level versus return period plot for (a) Tmin0305-Aut, (b) Tmin0305-Spr, (c) Tmin0305-Sum and (d) Tmin0305-Win (see online version for colours)

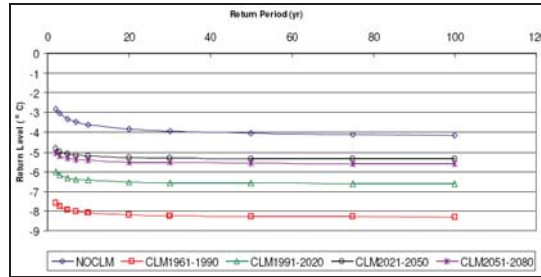


(a)

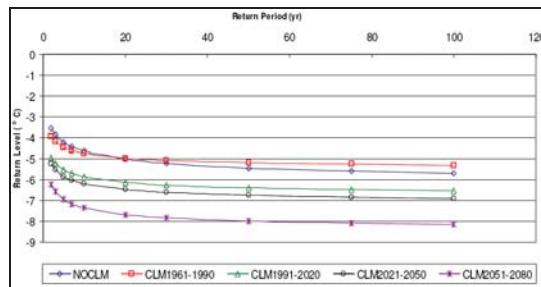


(b)

Figure 8 Return level versus return period plot for (a) Tmin0305-Aut, (b) Tmin0305-Spr, (c) Tmin0305-Sum and (d) Tmin0305-Win (continued) (see online version for colours)

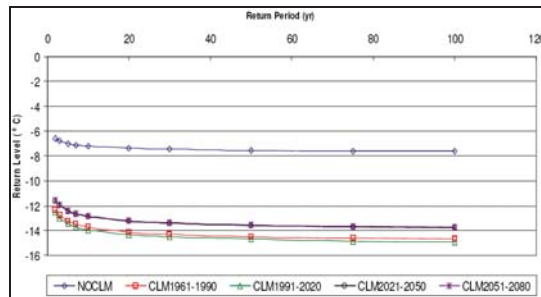


(c)

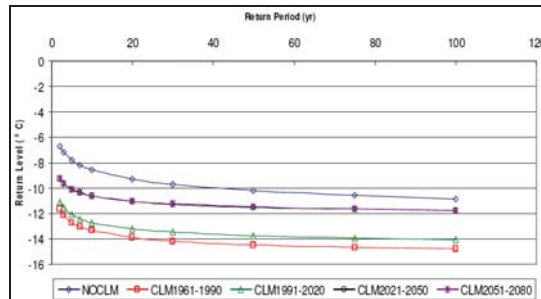


(d)

Figure 9 Return level versus return period plot for (a) Tmin4919-Aut, (b) Tmin4919-Spr, (c) Tmin4919-Sum and (d) Tmin4919-Win (see online version for colours)

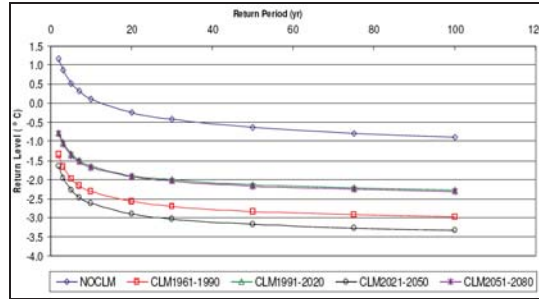


(a)

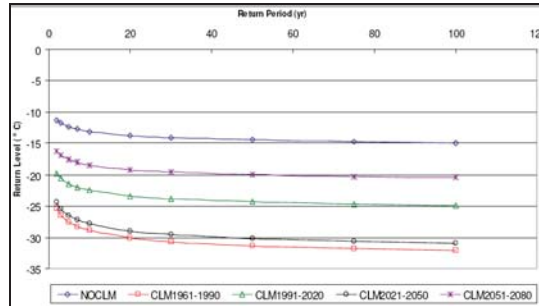


(b)

Figure 9 Return level versus return period plot for (a) Tmin4919-Aut, (b) Tmin4919-Spr, (c) Tmin4919-Sum and (d) Tmin4919-Win (continued) (see online version for colours)



(c)



(d)

For Birr station, the situation is somehow similar, and climate change influences appeared to be much stronger. The curves in Figures 9(a), 9(b), 9(c) and 9(d) suggest that a general decrease in Tmin is expected. The decrease is greater during the winter period, e.g., 15°C for a 100 years return period. The decrease is significantly reduced during the summer time, at 3°C. Accordingly colder conditions are expected at this location with this pattern of climate change. However, the extreme nature of the winter projections suggest a bias or error, possible associated with the parent GCM.

7 Summary and conclusions

In the present study, development of seasonal models for extreme maximum and minimum temperatures using the hybrid POT-GPD approach is addressed. The objective is to develop a seasonal model which is capable of maintaining the extreme value characteristics (magnitude and frequency) in addition to the ability to projecting climate change-related effects on these characteristics. The developed models are mainly based on fitting GPD to observed extreme events of both maximum and minimum temperatures, using a seasonally based POT theory as an extremal model. The reason for using this approach is that parameters of the fitted distribution are allowed to vary as function of some climatic variables (or covariates) selected from a large-scale atmospheric circulation pattern provided by a GCM model, at grid point level. Only the scale parameter is allowed to vary as function of the selected dominant covariates in the

named location. Observed data of maximum and minimum temperature from six stations, representing a mixture of both coastal and inland locations in Ireland, are used in building the seasonal models using the extReme software. Effective climate-driven return level-return period relations are also derived at each station based on the models developed. Results are presented in graphical and tabular formats and then analysed and discussed. Concluding remarks from this study can be summarised as follows:

- The hybrid POT-GPD seasonal models proposed here are proved to model the extreme behaviour of Tmax and Tmin in a very successful manner. The addition of covariates to predict future changes in distribution scale parameters has led to significant improvements in the model descriptive and predictive abilities.
- All developed seasonal models are suitable for use to obtain risky events of Tmax and Tmin quantiles for a given return period. All we needed are extraction of Tmax/Tmin POT series together with their covariates from NCEP data which has strongest relation 23 with them, generation of set of scale parameters using model parameters given in Tables 2 and 3 and corresponding GCM values of the associated covariates for the required future period, then calculate a series of possible values for the quantile. A maximum value for the Tmax quantile series (or a minimum for the Tmin quantile series) could then be selected to represent the desirable quantile.
- Tmax climate driven return level – return period relations derived in the study suggest that there is a possible increase in the number of extreme events for Tmax in Ireland between the current and future climate. The difference in temperature increase is in orders of 0.5 to 2°C, which is compatible with the GCM models projections of future increases in global temperature.
- Tmin climate driven return level – return period relations derived in the study suggest that there is a decrease likely in the number of extreme event of Tmin in Ireland with between the current and future time periods employed. This decrease, which seasonally varies, would translate into an expectation of milder winter conditions.
- The suggested increase in extreme temperature events, if realised, would have adverse effects on the natural environment, temperature-related public health and socioeconomic activities. Taking uncertainty in GCM model outputs, the proposed seasonal models and effective quantile return period relations developed here can only be used as tools during planning stage or as guidance for assessing or managing risk on environment-related activities affected by current global warming.

Acknowledgements

The authors would like to acknowledge the financial support provided by the Irish Environmental Protection Agency towards completion of this work as part of the Environmental RTDI Programme 2000–2006. Special thanks are due to Eric Gilleland and Richard Katz for providing the extRemes software and help with its installation and use. The authors also gratefully acknowledge Met Eireann, the Irish meteorological service, who provided the observed data for this study.

References

- Charles, S.P., Bates, B.C., Whetton, P.H. and Hughes, J.P. (1999) 'Validation of downscaling models for changed climate conditions: case study of southwestern Australia', *Climate Research*, Vol. 12, No. 1, pp.1–14.
- Coles, S. (2001) *An Introduction to Statistical Modelling of Extreme Values*, Springer Series in Statistics, London, UK.
- Cunnane, C. (1989) *Statistical Distributions for Frequency Analysis*, Operational Hydrology Report No. 33, Secretariat of the World Meteorological Organization, Geneva, Switzerland.
- Davison, A.C. (1984) 'Modelling excesses over high thresholds, with an application', in Tiago de Oliveira, J. (Ed.): *Statistical Extremes and Application*, pp.461–482, Reidel, Dordrecht.
- Demissie, S.S. (2004) *Effects of Climate Change on Rainfall Characteristics*, NUI Galway, Faculty of Engineering Research Day, Galway, Ireland.
- Frías, M.D., Mínguez, R., Gutiérrez, J.M. and Méndez, F.J. (2012) 'Future regional projections of extreme temperatures in Europe: a nonstationary seasonal approach', *Climate Change*, Vol. 113, No. 2, pp.371–392.
- Gilleland, E., Katz, R. and Young, G. (2005) *Extremes Toolkit (extRemes)*, Weather and Climate Applications of Extreme Value Statistics [online] <http://www.isse.ucar.edu/extremevalues/evtk.html> (accessed 05/03/2013).
- Henningsen, A. and Toomet, O. (2011) 'maxLik: a package for maximum likelihood estimation in R', *Comput. Stat.*, Vol. 26, pp.443–458.
- IPCC (2001) *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, 881p, Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K. and Johnson, C.A. (Eds.), Cambridge University Press, Cambridge, UK and New York, NY, USA.
- IPCC (2007) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, 996p, Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M. and Miller, H.L. (Eds.), Cambridge University Press, Cambridge, UK and New York, NY, USA.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, B., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R. and Joseph, D. (1996) 'The NCEP/NCAR 40-year reanalysis project', *Bulletin of the American Meteorological Society*, Vol. 77, No. 3, pp.437–472.
- Karl, T., Wang, W., Schlesinger, M., Knight, R. and Portman, D. (1990) 'A method for relating general circulation model simulated climate to the observed local climate. Part 1: seasonal statistics', *Journal of Climate*, Vol. 3, No. 10, pp.1053–1079.
- Katz, R.W. (1999) 'Extreme value theory for precipitation: sensitivity analysis for climate change', *Advances in Water Resources*, Vol. 23, No. 2, pp.133–139.
- Katz, R.W., Parlange, M.B. and Naveau, P. (2002) 'Statistics of extremes in hydrology', *Advances in Water Resources*, Vol. 25, No. 8, pp.1287–1304.
- Khaliq, M.N. and Cunnane, C. (1996) 'Modelling point rainfall occurrences with modified Bartlett-Lewis rectangular pulses model', *J. Hydrol.*, Vol. 180, No. 1, pp.109–138.
- Khaliq, M.N., Ouarda, T.B.M.J., Gachon, P. and Sushama, L. (2011) 'Stochastic modelling of hot weather spells and their characteristics', *Clim. Res.*, Vol. 47, No. 1, pp.187–199.
- Kiely, G. (1999) 'Climate change in Ireland from precipitation and stream flow observations', *Advances in Water Resources*, Vol. 23, No. 2, pp.141–151.
- Long, J.S. (1997) *Regression Models for Categorical and Limited Dependent Variables*, Sage, Thousand Oaks, CA.

- Milly, P.C.D., Betancourt, J., Falkenmark, M. et al. (2008) 'Stationarity is dead: Whither water management?', *Science*, Vol. 319, No. 5863, pp.573–574.
- Northrop, P.J. and Jonathan, P. (2011) 'Threshold modelling of spatially dependent nonstationary extremes with application to hurricane-induced wave heights', *Environmetrics*, Vol. 22, No. 7, pp.799–809.
- Palutikof, J.P., Holt, T., Brabson, B.B. and Lister, D.H. (2003) *Methods to Calculate Extremes in Climate Change*, School of Environmental Sciences, University of East Anglia, Norwich, UK.
- Rosenbrock, H.H. (1960) 'An automatic method of finding the greatest or the least value of a function', *Compu. J.*, Vol. 3, No. 3, pp.175–184.
- STARDEX Project (2003) *Statistical and Regional Dynamical Downscaling of Extremes for European Regions*, Deliverables D9, D10 and D11 [online] <http://www.cru.uea.ac.uk/projects/stardex> (accessed 05/03/2013).
- von Storch, H. and Zwiers, F.W. (1999) *Statistical Analysis in Climate Research*, Cambridge University Press, Cambridge, MA, USA.
- Wilby, R.L. and Dawson, C. (2007) *SDSM (Statistical Downscaling Model) Version 4.2* [online] <http://www-staff.lboro.ac.uk/~cocwd/SDSM/manual.html> (accessed 10/10/2012).
- Wilby, R.L. and Dawson, C.W. (2004) *Using SDSM Version 3.1 – A Decision Support Tool for the Assessment of Regional Climate Change Impacts*, User Manual.
- Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewiston, B.C., Main, J. and Wilks, D.S. (1998) 'Statistical downscaling of general circulation model outputs: a comparison of methods', *Water Resource Research*, Vol. 34, No. 11, pp.2995–3008.
- Wu, D. and Olson, D.L. (2009) 'Introduction to the special issue on 'risk issues in operation: methods and tools'', *Journal of Production Planning & Control*, Vol. 20, No. 4, pp.293–294.
- Wu, D. and Olson, D.L. (2010) 'Enterprise risk management: coping with model risk in a large bank', *Journal of the Operational Research Society*, Vol. 61, No. 2, pp.179–190.