

Statistical analysis and time-series models for minimum/maximum temperatures in the Antarctic Peninsula

BY GILLIAN L. HUGHES¹, SUHASINI SUBBA RAO²
AND TATA SUBBA RAO^{1,*}

¹*School of Mathematics, University of Manchester, Manchester M60 1QD, UK*
²*Texas A&M University, College Station, TX 77843-3143, USA*

It is now widely known that Antarctic air is warming faster than the rest of the world, and the Antarctic Peninsula has experienced major warming over the last 50 years. The monthly mean near surface temperature at the Faraday/Vernadsky station has increased considerably, at a rate of 0.56°C per decade over the year and at 1.09°C per decade over the winter. The increase is not the same over all the stations in the Antarctic region, and the increase is very significant at the Faraday/Vernadsky station. Only at this station are the minimum/maximum monthly temperatures, for the period 1951–2004, separately available, and we believe that the increase in mean surface temperature at this station is mainly due to the increases in minimum temperatures. Therefore, our object in this paper is to study the variations in the minimum/maximum temperatures using a multiple regression model with non-Gaussian correlated errors. By separately analysing the minimum and maximum temperatures, we could clearly identify the source of increase. The average temperature (usually calculated as $(\max + \min)/2$) smooths out any variation, and may not be that informative. We model the correlated errors using a linear autoregressive moving average model with innovations, which have an extreme value distribution. We describe the maximum-likelihood estimation methodology and apply this to the datasets described earlier. The methods proposed here can be widely used in other disciplines as well. Our analysis has shown that the increase in the minimum monthly temperatures is approximately 6.7°C over 53 years (1951–2003), whereas we did not find any significant change in the maximum temperature over the same period. We also establish a relationship between the minimum monthly temperatures and ozone levels, and use this model to obtain monthly forecasts for the year 2004 and compare it with the true values available up to December 2004.

Keywords: Antarctic Peninsula; extreme value distributions; ozone levels;
minimum/maximum temperatures; multiple regression; time-series models

1. Introduction

The data considered here are minimum/maximum monthly temperatures from the Faraday/Vernadsky research station in the Antarctic Peninsula from

* Author for correspondence (tata.subbarao@manchester.ac.uk).

January 1951 to December 2004. The dataset consists of 648 observations. The data were obtained from the British Antarctic Survey, Cambridge (and can be found on their website).

The daily average is calculated as the mean of the four, six hourly observations measured each day. The minimum/maximum temperature for each month for the early part of the record (1951–1986) was measured from a max and min thermometer that was reset everyday. Then it was logged on to a computer from 1986 onwards (information provided by Dr Steve Colwell, BAS, Cambridge, UK). For a complete description of the methods used to obtain the high-quality data and quality-control methods used, we refer to [Turner *et al.* \(2004\)](#).

The Faraday/Vernadsky data are of interest because this is the longest continuous record of any British Antarctic station. The whole Antarctic Peninsula has experienced a major warming over the last 50 years, with major warming of mean monthly near surface temperatures at the Faraday/Vernadsky station ([Houghton *et al.* 2001](#); [Turner *et al.* 2002, 2005](#)). It has been observed that over the past 50 years, the annual mean temperatures at this station have risen by approximately 1.09°C per decade over the winter, and above 0.59°C per decade over the year. Although the temperature increase at the Faraday/Vernadsky is substantial, there are other indications that the region of marked warming is quite limited and restricted to an arc from the southern part of the Peninsula through Faraday/Vernadsky to a little beyond the tip of Peninsula.

[Turner *et al.* \(2005\)](#) have fitted linear trend models to mean monthly surface temperatures at 19 selected Antarctic stations from the datasets available from each station. Eleven of these had warming trends and seven had cooling trends (one station had very little data and hence was omitted for their analysis), though the most statistically significant increase was found at the Faraday/Vernadsky station. The Faraday/Vernadsky station is of interest to us for two reasons, firstly because it has the longest continuous record compared to any other station, and secondly because it has both mean monthly temperatures and also minimum/maximum monthly temperatures. This allows us to identify the main source of increase at this station. Unfortunately, due to the lack of availability of minimum/maximum temperatures at other stations, we cannot perform our analysis for other stations, though this analysis would be possible if such minimum/maximum temperatures were made available. We believe the increase at Faraday may be due to an increase in minimum temperatures rather than in the maximum temperatures. Turner and his co-authors have fitted linear trend models to annual temperatures, and also to the four seasons, spring (September, October and November), summer (December, January and February), autumn (March, April and May) and winter (June, July and August) for the years January 1951–December 2005, thus using 55 observations for each analysis. The measurements used by [Turner *et al.* \(2005\)](#) are obtained by averaging the monthly data, and by doing this they have smoothed the data, thereby removing any variation in the months considered. Plots of the annual surface temperatures and the seasonal surface temperatures at the Faraday station in [Turner *et al.* \(2005\)](#) clearly show a linear upward trend. The most significant upward trend at the Faraday/Vernadsky station is observed during the winter months. They computed the linear trend for mean surface temperatures using ordinary least squares (OLS) and the confidence limits were calculated under

the assumption that the random errors were independent and identically normally distributed. When we analysed this data, we observed that the residuals obtained (after removing the trend) were strongly correlated at several lags and also the residuals were highly skewed indicating a departure from Gaussianity. This suggests that the data require further statistical analysis which takes into account the correlation and departure from Gaussianity.

Possible reasons for causes in warming at Faraday/Vernadsky were considered in King (1994) and King & Harangozo (1998). Using the climatological records available at several stations in the western Peninsula, King & Harangozo (1998) observed considerable interannual variability and identified two major factors for the variability. First is the atmosphere–ice–ocean interaction, which is seen to play an important role in controlling variability in the Peninsula. There seems to be a strong negative correlation between winter temperatures in the region and the winter sea-ice content just to the west of Peninsula. The Peninsula winter temperatures are strongly correlated with the atmospheric flow and circulation (for details refer to King & Harangozo 1998). Second is the variability in the advection of warm masses which exerts an important control on climate (for details see King & Harangozo 1998).

We believe this interesting relationship found between warming at Faraday/Vernadsky and other variables studied by King and Harangozo requires further detailed statistical analysis. A consequence of warming over the winter months is that the small fringing of ice shelves around the Antarctic Peninsula is retreating (Hulbe 1997; Turner *et al.* 2005). The change in local climate is also demonstrated by profound ecological changes (Smith *et al.* 2003).

An increase in atmospheric temperature, if large enough to push summer temperatures above freezing point, will increase mass loss directly by increasing melting at the upper surface. Warmer sea surface temperatures may accompany warmer air temperatures and this could also increase the rate of ice shelf melting. There are also indirect effects of warmer air temperatures that can hasten the decay of ice shelves.

The effects of CO₂-induced climate warming on Antarctica have been studied using numerical models that simulate ice flow and changes in ice sheet and ice shelf size over time (Huybrechts & Oerlemans 1990; Budd *et al.* 1994). One prediction of the models is that the glaciers and ice shelves of the Peninsula are lost.

An increase in minimum temperatures could have a more dramatic effect on the surroundings than an overall increase in temperatures. There is clearly a connection between warming around the Antarctic Peninsula and the collapse of Peninsula ice shelves, and unless there is a change in the observed warming trend, further retreats of fringing of ice shelves along the Antarctic Peninsula are inevitable. It is therefore important to forecast how quickly the temperatures are going to increase above melting point as this is vital to the break-up of ice shelves. Clearly, if the minimum summer temperatures increase above 0°C, the melt season will be tremendously long and the ice shelves are likely to become very unstable. Our study attempts to analyse the trend in the minimum monthly temperatures. We consider both maximum and minimum monthly temperatures, though our main emphasis is on minimum temperatures and we also consider its relationship to ozone levels.

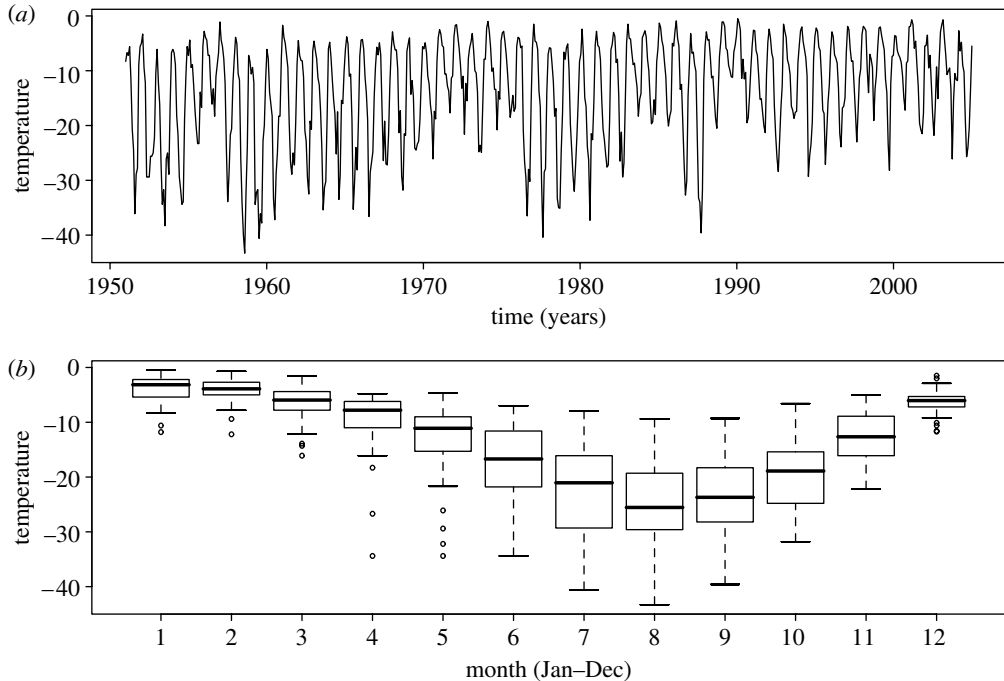


Figure 1. (a) Minimum monthly temperatures and (b) boxplot of minimum monthly temperatures at the Faraday station (January 1951–December 2004).

2. Preliminary data analysis

(a) Data analysis of minimum monthly temperatures

In figure 1, the minimum monthly temperature series from January 1951 to December 2004 is plotted. From the plots of the monthly mean Antarctic surface temperatures during the aforementioned period (Turner *et al.* 2005), it appears that summer temperatures after 1950 are more often above 0°C than before. The data we plotted here are minimum monthly temperatures and we are examining whether there is a significant increase in minimum monthly temperatures. One of our main aims in this paper is to estimate the trend using appropriate statistical time-series models and to see whether the coefficients are significantly different from zero. The boxplot of the minimum monthly temperature data by month, where month 1 is January, month 2 is February and so on is also given in figure 1. Clearly there is a yearly cycle, but it is also seen that there is a much larger spread of temperatures in the winter months, i.e. from June to August. The range of the entire series is (-43.3°C , -0.5°C). There is some indication that the variance is changing with the season. During the winter months, there are no ‘extreme values’ lying outside of the ‘whisker’ on the boxplots. The whisker shown is 1.5 times the inter-quartile range. Extreme values (outliers) only exist for the months February to May. Furthermore, these are only on the lower temperature side, implying that the extreme temperatures are lower and not higher than the median. There is a clear increase in the yearly median value over time. We consider the data analysis of ozone levels at the same station and its relationship with temperature in a later section.

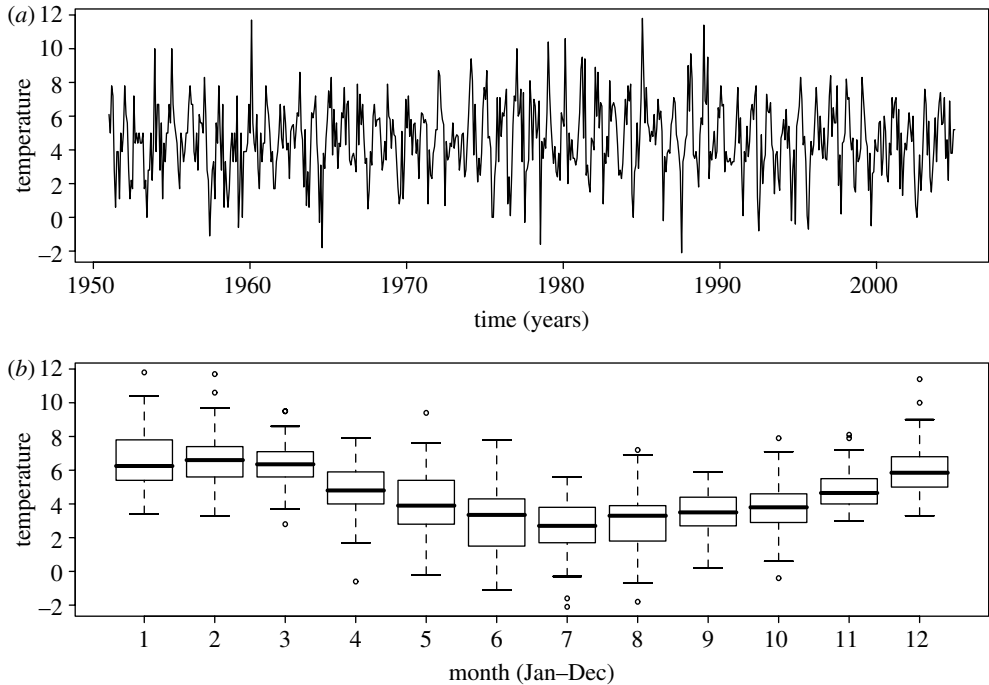


Figure 2. (a) Maximum monthly temperatures and (b) the boxplot of maximum monthly temperatures at the Faraday station (January 1951–December 2004).

(b) *Data analysis of maximum monthly temperatures*

We now analyse the maximum monthly temperatures to see whether there is any asymmetric climate warming at the Faraday station. A plot of the maximum monthly temperature data and boxplot by month are given in figure 2. The range of this series (-2.1°C , 11.8°C) is quite small compared to the minimum monthly temperatures. There is not a significant change over the period 1951–2004 compared to the minimum temperatures. From the boxplot given in figure 2, we see that there is possibly a yearly cycle, and in most of the months some extreme values have occurred. The median and range of values over the 54 years seems to have remained more or less the same compared to the minimum temperatures.

(c) *Multiple regression of minimum temperature using ordinary least squares*

From the aforementioned preliminary data analysis, we see that minimum temperatures are exhibiting more dynamic behaviour. As far as we are aware, no attempts have been made to analyse, by the regression methods, the minimum/maximum temperatures and all the analyses we have seen so far are for monthly, annual and seasonal data.

To motivate the analysis below, we fit a multiple regression model of the form given in equation (4.1) to the minimum monthly temperatures with a linear trend and seasonal component with a known periodicity of 12 months. We use OLS to estimate the parameters in the regression and give a normal $Q-Q$ plot and autocorrelation plot of the residuals (removing the trend) in figure 3.

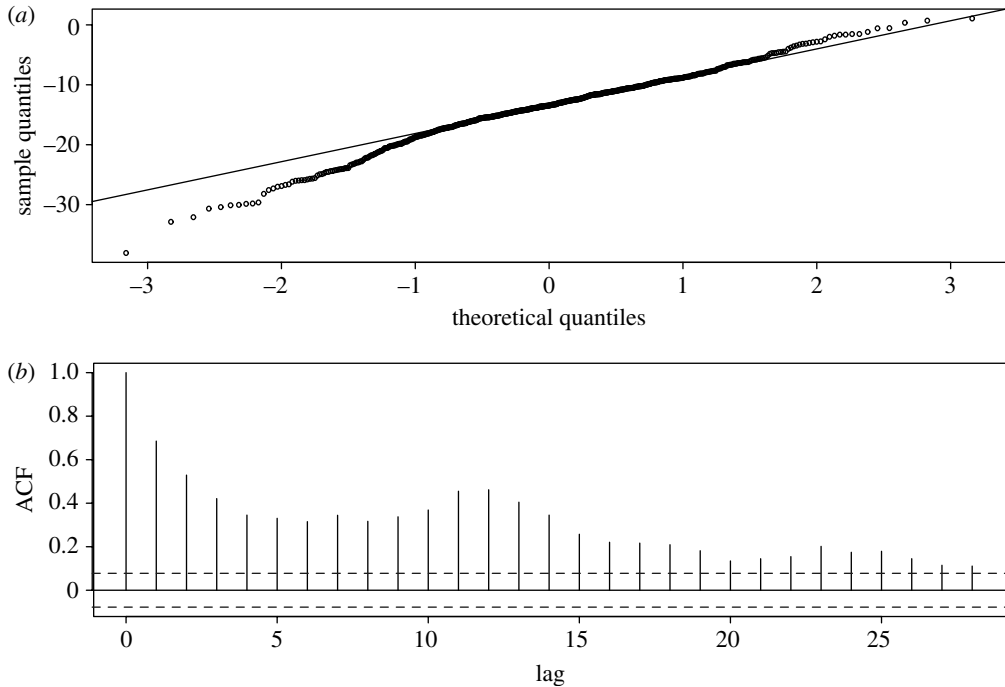


Figure 3. (a) Normal $Q-Q$ plot of the residuals and (b) autocorrelation plot of the residuals.

The $Q-Q$ plot is a graphical technique for assessing whether a dataset follows a given distribution. Here, we are checking for the validity of the normal assumption. There appears to be two striking features in [figure 3](#): (i) we see in the $Q-Q$ plot that *ca* 35% of the residuals lie far from the $x=y$ line, in fact an estimate of the kurtosis, $\left(\frac{(1/n) \sum_{i=1}^n \hat{\varepsilon}_i^4}{\left(\frac{(1/n) \sum_{i=1}^n \hat{\varepsilon}_i^2}{n}\right)^2}\right)$ yields 3.85, which is sufficiently different from three for us to conclude that the temperature residuals are highly non-normal; and (ii) the temperature residuals are highly correlated. Now it is well known ([Chatfield 2004](#), pp. 87–88) that if multiple regression models are fitted ignoring the correlation in errors, the models can lead to badly misspecified models and poor forecasts, and to estimates of the regression parameters with large mean square errors and wide confidence intervals. All these can lead to wrong statistical conclusions.

In the following sections, we briefly discuss the estimation methodology for minimum and maximum temperatures. We consider the estimation of multiple regression models where the errors are correlated and the errors can be described by linear autoregressive moving average (ARMA) models with innovations, which have extreme value distributions. The methodology and the steps of estimation are very general and are applicable in many other situations.

3. Multiple regression with correlated errors

In this section, we consider the maximum-likelihood estimation of the parameters of a classical multiple regression model with correlated errors, when the errors satisfy the ARMA model with innovations having a generalized

extreme value (GEV) distribution (see appendix A). Suppose we observe the time-series $\{y_t\}$, which satisfies

$$y_t = \sum_{j=1}^r \beta_j x_{j,t} + e_t; \quad t = 1, \dots, n, \quad (3.1)$$

and $\{e_t\}$ satisfies the stationary ARMA (p, q) model

$$(1 + \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p) e_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \eta_t$$

or equivalently, $\phi_p(B) e_t = \theta_q(B) \eta_t,$ (3.2)

where $\{\eta_t\}$ are independent, identically distributed random variables and each η_t has the converse GEV (γ, μ, σ) distribution. In the above, we have used the notation $Bx_t = x_{t-1}$, where B is a backward shift operator and $\phi_p(B)$ and $\theta_q(B)$ are p and q order polynomials, respectively. We assume that the regressor variables $(x_{j,t})$ are non-random. In many real situations, the regressors are known *a priori* and the number of variables (r) to be used is also known. However, we note that there are various statistical procedures available for the choice of the best set of variables, which we do not discuss in this paper. The time-series $\{e_t\}$ generated by equation (3.2) is assumed to be stationary and the model is invertible. Using equation (3.2), we can write equation (3.1) as

$$y_t = \sum_{j=1}^r \beta_j x_{j,t} + \phi_p(B)^{-1} \theta_q(B) \eta_t, \quad (3.3)$$

noting that $\phi_p(B)^{-1}$ is an infinite order polynomial expansion in terms of B (Chatfield 2004, p. 46), or equivalently,

$$\theta_q(B) \eta_t = \phi_p(B) \left[y_t - \sum_{j=1}^r \beta_j x_{j,t} \right]; \quad (t = 1, 2, \dots, n). \quad (3.4)$$

Let $s = \max(p, q)$. Assuming that the probability density function (PDF) of η_t is a converse GEV distribution (denoted by conv GEV (γ, μ, σ) and defined in equation (A 3) in appendix A), we can write the conditional log-likelihood function of $(\eta_{s+1}, \eta_{s+2}, \dots, \eta_n)$ as

$$\begin{aligned} \mathfrak{l}(\eta; \underline{\beta}, \underline{\phi}, \underline{\theta}, \gamma, \mu, \sigma) &= -(n-s) \log \sigma - \left(\frac{1}{\gamma} + 1 \right) \sum_{t=s+1}^n \log \left[1 + \gamma \left(\frac{\mu - \eta_t}{\sigma} \right) \right] \\ &\quad - \sum_{t=s+1}^n \left[1 + \gamma \left(\frac{\mu - \eta_t}{\sigma} \right) \right]^{-1/\gamma}, \end{aligned} \quad (3.5)$$

provided $1 + \gamma((\mu - \eta_t)/\sigma) > 0$ for each t , where $\underline{\beta} = (\beta_1, \dots, \beta_r)$, $\underline{\phi} = (\phi_1, \phi_2, \dots, \phi_p)$, $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_q)$.

By using equations (3.1) and (3.2), we have

$$\eta_t = \sum_{i=0}^p \phi_i \left(y_{t-i} - \sum_{j=1}^r \beta_j x_{j,t-i} \right) - \sum_{k=1}^q \theta_k \eta_{t-k}, \quad (3.6)$$

where $\phi_0 = 1$. In order to estimate the parameters $\underline{\beta}, \underline{\phi}, \underline{\theta}, \gamma, \mu, \sigma$, we substitute the aforementioned expression for η_t into equation (3.5) and maximize $\ell(\cdot)$ with respect to the parameters of interest. However, we note that $\{\eta_{t-k}\}$ on the right-hand side of equation (3.6) is unobserved and needs to be estimated in order to evaluate equation (3.5). This can be done using equation (3.4) and the following recursive procedure, since $\theta_q(B)$ and $\theta_p(B)$ are finite-order polynomials (Chatfield 2004, pp. 46–47) by setting $\eta_t \equiv 0$ and $y_t \equiv 0$ for all $t \leq 0$, we can estimate $\{\eta_t; t=1, \dots\}$ using initial estimates of the multiple regression parameters and the ARMA parameters (obtained following the steps mentioned later). In other words, we define $\{\hat{\eta}_t\}$ using the recursion given in equation (3.7). We mention that an invertible ARMA process can be represented as an $\text{AR}(\infty)$ process, whose parameters decay exponentially, and $\hat{\eta}_t$ is the $\text{AR}(\infty)$ representation truncated at lag t ; therefore, setting $\eta_t \equiv 0$ and $y_t \equiv 0$ for all $t \leq 0$ means that $\hat{\eta}_t$ will get closer to η_t as t increases.

The maximization of the log-likelihood function (3.5) is done using a Newton–Raphson iterative procedure, using the OLS estimates as our initial estimates and replacing $\{\eta_{t-k}\}$ in equation (3.6) by $\{\hat{\eta}_{t-k}\}$, and doing the maximum-likelihood estimation as if $\{\eta_{t-k}\}$ were known. By choosing the OLS estimates as our initial values, we are sufficiently close to the maximum of the likelihood for the algorithm to converge to the maximum. The first and the second order derivatives of the likelihood can be obtained either numerically or analytically and the analytical expressions can be found in Hughes (2002).

In the following, we describe the steps for obtaining the initial estimates of the parameters. We note that the initial estimates are consistent estimators of the parameters, but are less efficient compared to the maximum-likelihood estimators described earlier.

Initial parameter estimators:

- (i) First estimate the regression parameters, $\underline{\beta}$, by the method of OLS. Let $\hat{\underline{\beta}}$ be such an estimate.
- (ii) Obtain

$$\hat{e}_t = y_t - \sum_{j=1}^r \hat{\beta}_j x_{j,t}; \quad (t = 1, 2, \dots, n).$$

Test the residuals $\{\hat{e}_t\}$ for zero correlation (cf. Brockwell & Davis 1996). If we reject the null hypothesis, we fit an $\text{ARMA}(p, q)$ model to $\{\hat{e}_t\}$ using the Hannan–Rissanen (1982) method. The orders p and q are chosen using the Bayes information criterion (BIC; for further details on the use of the Akaike information criterion and the BIC we refer to Davison 2003). Let \hat{p} and \hat{q} be the chosen orders and $\hat{\underline{\phi}}$ and $\hat{\underline{\theta}}$ be the Hannan–Rissanen estimators of $\underline{\phi}$ and $\underline{\theta}$.

- (iii) Using equation (3.4) and the earlier discussion, we obtain the estimated residuals

$$\hat{\eta}_t = \sum_{i=0}^{\hat{p}} \hat{\phi}_i \left[y_{t-i} - \sum_{j=1}^r \hat{\beta}_j x_{j,t-i} \right] - \sum_{k=1}^{\hat{q}} \hat{\theta}_k \hat{\eta}_{t-k}, \quad (3.7)$$

where $\hat{\phi}_0 = 1$, $\hat{\eta}_t \equiv 0$ and $y_t \equiv 0$ for $t \leq 0$. The residuals, $\{\hat{\eta}_t\}$, are tested for Gaussianity using the standard skewness and kurtosis measures (D’Agostino & Stephens 1998). If we reject the null hypothesis, we have to search for an

appropriate distribution. Since we are considering in this paper the minimum and maximum temperatures, a natural family of distributions are extreme value distributions. The parameters, γ , μ and σ , of the converse GEV distribution are estimated using the computer package ISMEV (found in the R library).

It is useful to test the residuals using the probability plots for checking the validity of the assumption on the distribution.

The asymptotic properties of the maximum-likelihood estimates of the parameters of the converse GEV distribution (γ, μ, σ) , as given in the appendix, were studied in [Smith \(1985\)](#) when one has a random sample from the above distribution. His conclusion was that if the shape parameter, γ , lies between $-\frac{1}{2} < \gamma < \infty$, the Fisher information exists and if $-\frac{1}{2} < \gamma < 0$, then all the asymptotic properties of the maximum-likelihood estimates (such as consistency and normality) still hold. However, we must note that we do not observe η_t , and these have been estimated after fitting a multiple regression model. Whether the results of [Smith \(1985\)](#) still hold in the present context needs to be studied and is beyond the scope of the present paper. This is an interesting and quite challenging problem, which we hope to consider in the future. From our estimation of the shape parameter, using the methods described earlier for the minimum/maximum temperatures considered, we believe that the results of [Smith \(1985\)](#) still hold. The simulation studies carried out by [Hughes \(2002\)](#) confirm this to some extent.

4. Time-series model for the temperatures

(a) Model for the minimum temperatures

The aforementioned methodology also holds when the regressors are of the form $x_{j,t} = t^j$, $x_{j,t} = \sin(\omega_j t)$ or $x_{j,t} = \cos(\omega_j t)$ or a mixture of both polynomials and harmonic terms. For the temperature data considered here, there is clear evidence that there is a dominant 12-month cycle and hence the problem of estimating frequencies does not arise. If these were unknown, they can be estimated using the methods advocated in [Kavalieris & Hannan \(1994\)](#) and also in [Quinn & Fernandes \(1991\)](#).

Following a detailed estimation and order determination procedure, we concluded that the minimum monthly temperatures could be described by a linear trend, one periodic term (with a 12-month period) and an error term satisfying a first-order autoregressive model with innovations satisfying a conv GEV(γ, μ, σ) distribution. We use the procedure described in §3 to estimate the parameters; however, due to space limitations, we do not give specific details, but instead, refer to [Hughes \(2002\)](#).

We also compared this model with other models where we included/excluded the trend, independent Gaussian/dependent Gaussian and independent converse GEV/dependent converse GEV. We summarize our analysis in [table 1](#), where the linear trend is excluded and [table 2](#), where the linear trend is included. We note that the bracket values are the estimated standard errors of the corresponding parameter estimators, where the standard errors were calculated using an estimate of the Fisher information matrix. Let $\{y_t\}$ be the monthly minimum

Table 1. Minimum temperature data from 1951 to 2003, modelled using a seasonal term.

parameter/s.e.	ζ	A	B	ϕ	μ	σ	γ	BIC	prediction MSE
independent converse GEV		5.74 (0.370)	5.56 (0.301)		-11.1 (0.222)	4.65 (0.165)	-0.0048 (0.0377)	6.22	12.7
independent Gaussian	-18.1 (0.406)	7.70 (0.286)	7.28 (0.0011)					6.10	40.1
dependent converse GEV		5.87 (0.471)	7.17 (0.392)	0.625 (0.028)	-3.3 (0.406)	3.37 (0.114)	-0.103 (0.024)	5.65	12.7
dependent Gaussian	-13.8 (0.438)	7.66 (0.495)	7.26 (0.438)	0.685 (0.028)				5.69	11.12

Table 2. Minimum temperature data from 1951 to 2003 modelled with a seasonal term and linear trend.

parameter/s.e.	ζ	A	B	β	ϕ	μ	σ	γ	BIC	prediction MSE
independent GEV		6.47 (0.341)	5.85 (0.303)	0.0110 (0.00107)		-15.0 (0.431)	4.40 (0.153)	-0.058 (0.035)	6.028	16.47
independent Gaussian	-18.1 (0.406)	7.70 (0.286)	7.28 (0.0011)	0.013 (0.0011)					6.01	22.9
dependent GEV		6.25 (0.441)	6.95 (0.368)	0.0105 (0.00193)	0.566 (0.03)	-5.71 (0.602)	3.65 (0.114)	-0.109 (0.024)	5.62	11.09
dependent Gaussian	-18.1 (0.81)	7.70 (0.495)	7.26 (0.404)	0.0136 (0.0022)	0.609 (0.031)				5.65	11.20

temperatures. The models we fitted were of the form

$$y_t = \zeta + A \sin\left(\frac{2\pi}{12}t\right) + B \cos\left(\frac{2\pi}{12}t\right) + \beta t + \varepsilon_t, \quad (4.1)$$

where the errors $\{\varepsilon_t\}_t$: (i) are iid converse GEV(γ, μ, σ); (ii) are iid Gaussian; (iii) satisfy $\varepsilon_t = \phi\varepsilon_{t-1} + \eta_t$ with $\{\eta_t\}$ iid Gaussian; and (iv) satisfy $\varepsilon_t = \phi\varepsilon_{t-1} + \eta_t$ with $\{\eta_t\}$ iid converse GEV(γ, μ, σ). We mention that when we are using the converse GEV distribution, we set the intercept parameter ζ to 0 to avoid over parameterization, as the mean is absorbed into the mean of the converse GEV distribution. In both tables, the BIC values are given; we note that typically the model with the smallest BIC is chosen. All the models are fitted using the monthly temperature data from January 1951 to December 2003, the resulting model is used to obtain monthly forecasts for January–December 2004. The estimated mean square error of the one-step ahead prediction is also given in tables 1 and 2. Exact details on how the forecasts and mean squared errors are calculated are given below.

On examining tables 1 and 2, we see that the linear trend β is highly significant. We also note that there is a strong correlation (see the coefficients ϕ) in the errors, which is an important factor in forecasting. The most optimal model on the bases of the minimum BIC and prediction mean squared error seems to be the following:

$$y_t = 0.0105t + 6.25 \sin\left(\frac{2\pi}{12}t\right) + 6.95 \cos\left(\frac{2\pi}{12}t\right) + e_t,$$

$$e_t = 0.566e_{t-1} + \eta_t, \quad \text{where } \eta_t \sim \text{converse GEV}(-0.109, -5.71, 3.65). \quad (4.2)$$

(b) Goodness of fit test by probability plots

We now confirm the hypothesis that the residuals $\{\hat{\eta}_t\}$ follow a converse GEV distribution, by using the probability plot. Suppose the random variable X has a converse GEV distribution with parameters (γ, μ, σ) , then if

$$p_t = P(X \leq x_t) = 1 - \exp\left\{-\left[1 + \frac{\gamma}{\sigma}(\mu - x_t)\right]^{-1/\gamma}\right\},$$

this implies

$$-\log(1 - p_t) = \left[1 + \frac{\gamma}{\sigma}(\mu - x_t)\right]^{-1/\gamma}.$$

In addition, the plot $[1 + (\gamma/\sigma)(\mu - x_t)]^{-1/\gamma}$ against $-\log(1 - p_t)$ is a validation plot. Let $(\eta_1, \eta_2, \dots, \eta_n)$ be a random sample from a converse GEV(γ, μ, σ), and let $(\eta_{(1)}, \eta_{(2)}, \dots, \eta_{(n)})$, be its order statistics. Let $p_t = ((t - 0.5)/n)$, $t = 1, 2, \dots, n$. In figure 4, we plot $y_t = [1 + (\hat{\gamma}/\hat{\sigma})(\hat{\mu} - \eta_{(t)})]^{-1/\gamma}$ against $x_t = -\log(1 - p_t)$, where $\hat{\gamma}, \hat{\mu}, \hat{\sigma}$ are the maximum-likelihood estimates and $\hat{\eta}_t$ is the estimate of η_t . We observe that the majority of the observations (y_t, x_t) fall on the $y = x$ line (which passes through the origin). Only 18 points lie significantly off the $y = x$ line, this is less than 3% of the total number of observations. Compare this with the $Q-Q$

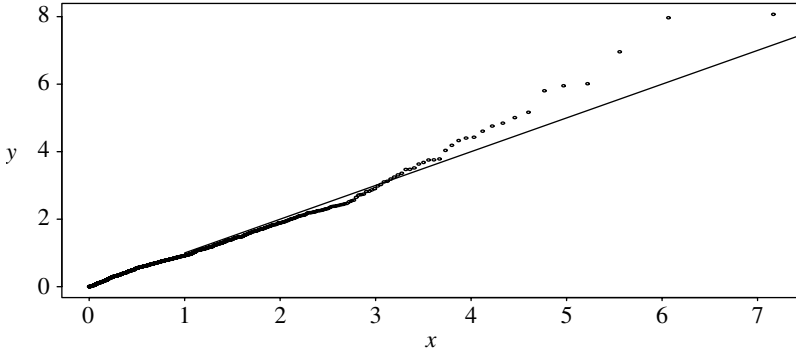


Figure 4. Validation plot: $y_t = [1 + (\hat{\gamma}/\hat{\sigma})(\hat{\mu} - \eta_{(t)})]^{-1/\gamma}$ against $x_t = -\log(1 - p_t)$.

plot in figure 3, where *ca* 35% of the points lie significantly off the $y=x$ line. This justifies the use of the converse GEV distribution, which confirms that our assumptions on the distribution are appropriate.

(c) *Prediction of the minimum temperatures*

Consider the time-series $\{y_t\}$ where y_t satisfies the model (4.2). Suppose we have observations $\{y_t, s \leq t_0\}$, then the minimum mean square error forecast of y_{t_0+l} ($l=1, 2, \dots$) is given by

$$\begin{aligned} y_{t_0}(l) &= E(y_{t_0+l} | y_s, s \leq t_0) \\ &= \beta(t_0 + l) + A \cos(0.523(t_0 + l)) + B \sin(0.523(t_0 + l)) + e_{t_0}(l), \end{aligned}$$

where $e_{t_0}(l) = \phi e_{t_0}(l-1) + E(\eta_t)$

$$E(\eta_t) = \mu + \frac{\sigma}{\gamma} \{1 - \Gamma(1 - \gamma)\} \quad (4.3)$$

(see appendix A). Thus, when $l=1$, $E(y_{t_0+1} - y_{t_0}(1))^2 = \text{var}(\eta_t) = \sigma^2$.

Let $\hat{y}_{t_0}(1)$ be the one-step ahead forecast, where β , A , B , γ , μ and σ in equation (4.3) have been replaced with their maximum-likelihood estimates. In figure 5, we plot the one-step ahead forecasts $\hat{y}_{t_0}(1)$, for the period January–December 2004. The estimated mean square error $\hat{\sigma}_{\eta}^2 = (1/12) \sum_{t=636}^{647} (\hat{y}_t(1) - y_{t+1})^2$ of the one-step ahead forecast over this period is $\hat{\sigma}_{\eta}^2 = 11.09$. There is a close agreement between the forecasts and the true values.

(d) *Model for the maximum temperatures*

We now consider the maximum monthly temperatures for the period January 1951–December 2003.

As pointed out earlier, the range of the series is (-2.1°C , 11.8°C), which is quite small compared to the minimum monthly temperatures. From the plot of the series, we cannot see any significant trend in the series. There is a yearly cycle, and from the boxplot (figure 2), we see that there are extremes in the

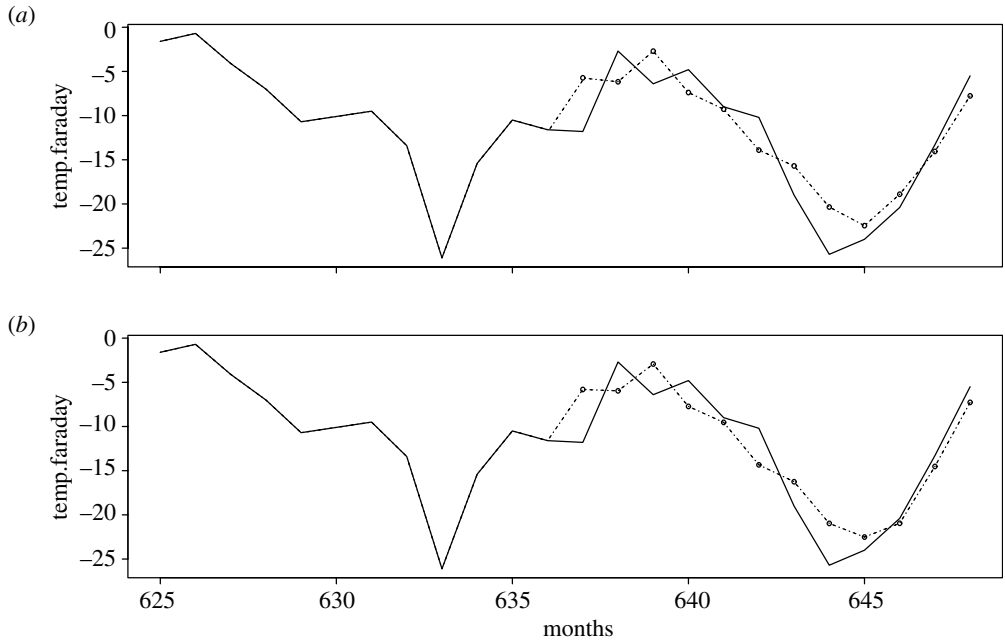


Figure 5. (a) One-step ahead forecasts of temperatures $\{y_t\}$ from January to December 2004 using model (4.2). (b) One-step ahead forecasts of temperatures $\{y_t\}$ from January to December 2004 using model (5.2). Both plots are given from January 2003 to December 2004. The dotted lines are the one-step ahead predictors and thick lines are the true values.

summer months (January, February and March) and in the winter months (June, July and August). The extremes in summer months may cause melting of the ice shelves. In view of these observations, we do not include a trend component in the model. Since we are analysing the maxima, it is natural for us to assume that the innovations have a GEV distribution, $\text{GEV}(\gamma, \mu, \sigma)$. We use the same procedure for estimating the model as before. The time-series model for the maximum temperature is found to be

$$y_t = \underset{(0.0003)}{0.0002t} + \underset{(0.095)}{1.74} \cos(0.523t) + \underset{(0.095)}{1.08} \sin(0.523t) + e_t, \quad (4.4)$$

$$e_t \sim \text{GEV}(\underset{(0.0004)}{-0.260}, \underset{(0.095)}{1.08}, \underset{(0.047)}{1.67}).$$

We note that there is no significant trend. It is interesting to note that the errors $\{e_t\}$ are mutually independent, but non-Gaussian. This is in contrast to the minimum temperature data, which has a significant linear trend and correlated errors. This implies that the diurnal temperature range is decreasing and that asymmetric climate change is occurring in the Antarctic Peninsula.

Since the minimum temperatures are increasing and the maximum temperatures are remaining constant over the same period, we believe that the changes in minimum temperatures are more significant. In §5, we investigate this data further and see whether there is any relationship between minimum temperatures and ozone levels at the Faraday/Vernadsky station. The ozone meteorological and ozone-monitoring unit, BAS, issues regular ozone bulletins describing ozone levels at various stations in the Antarctic. In recent years, when

there was an increase in stratospheric temperatures, generally there was a decrease in ozone levels at the Halley and Faraday/Vernadsky stations. For our analysis, we consider time-series analysis of minimum temperatures and mean monthly ozone levels at the Faraday/Vernadsky station for the period January 1958–December 2004, giving us 564 observations. (This data was made available to us by Dr Jon Shanklin, Dr S. Colwell and Dr John Turner, BAS.)

5. Effect of mean monthly ozone levels on minimum monthly temperatures

The influence of the human race on climate is still a matter for study and speculation, but the ability to perturb the ozone layer is an established fact. We next examine if the amount of ozone in the stratosphere in the Antarctic Peninsula has a direct relationship to the minimum temperatures. If this can be established, then it can be deduced that human activity does play some role in increasing the temperatures in the Antarctic Peninsula and the future temperatures can be predicted with more certainty. These stratospheric ozone concentrations are recorded in Dobson units using a Dobson ozone spectrophotometer. This instrument tells us how much ozone there is in the atmosphere by comparing the intensities of two wavelengths of ultraviolet light from the Sun. It is therefore not possible to make regular measurements of ozone during the Antarctic winter because the station is in darkness. These missing values have been substituted by their yearly average.

Comparing the plots of the minimum temperatures and the ozone levels in figures 1 and 6, we see that there is a decrease in ozone levels during the years 1980–2001 and during the same period there is a steady increase in the minimum temperatures $\{y_t\}$.

For our cross-correlation analysis, we consider the detrended deseasonalized minimum monthly temperatures and similarly the ozone levels which have also been pre-whitened (e.g. Chatfield 2004, p. 158) during the period January 1958–December 2004. The plot of the cross correlations for various lags is given in figure 6. There is a strong negative correlation at all lags. Further, we see that there is a significant negative correlation at lag (-1) indicating that ozone levels affect the future minimum temperature levels (at least one month later). Our analysis also indicates that the ozone levels may have a long-term effect, but the data we have is not sufficiently long to draw any significant conclusions.

We now consider the regression of the minimum temperatures $\{y_t\}$ on the detrended ozone $\{x'_{t-1}\}$ (where seasonal and linear trend in the ozone has been removed). We have detrended the ozone to prevent a cointegration effect between temperature and ozone (cointegration can loosely be described as spurious correlation between two variables caused by a common trend). We fit models of the form

$$y_t = \zeta + A \sin\left(\frac{2\pi}{12}t\right) + B \cos\left(\frac{2\pi}{12}t\right) + \alpha x'_{t-1} + \beta t + e_t, \quad (5.1)$$

where $\{e_t\}$ are random errors. Using the procedure described in §3 to estimate the parameters, we concluded that the minimum monthly temperatures can be described by a linear trend, one periodic term (with a 12-month period), the

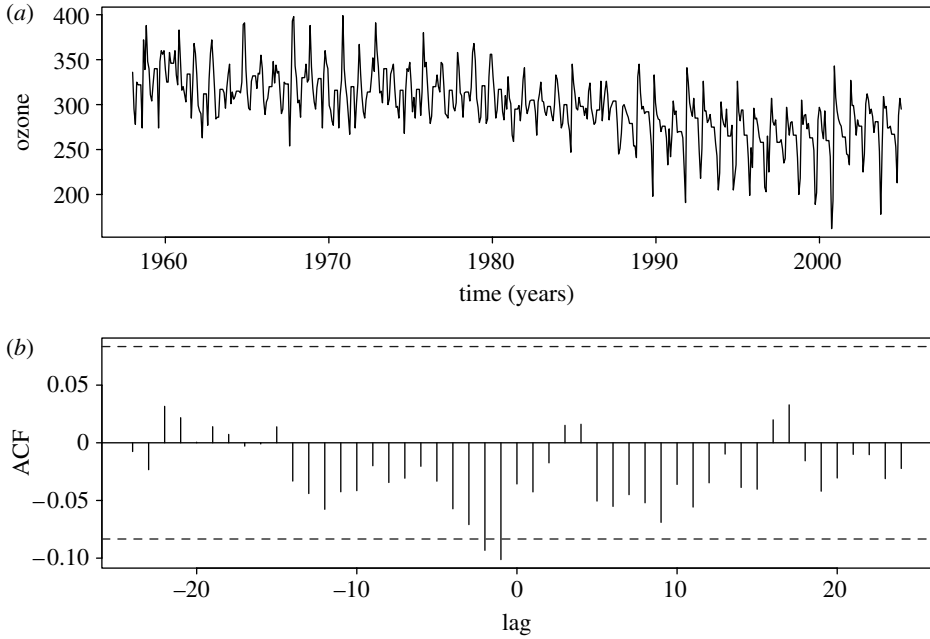


Figure 6. (a) Ozone levels at the Faraday station (1958–2005) and (b) the cross correlation between temperature and pre-whitened ozone.

detrended ozone at lag (-1) and an error term satisfying a first-order autoregressive model with innovations satisfying a conv $GEV(\gamma, \mu, \sigma)$ distribution. Using the ozone at additional lags is not significant and hence we have excluded them from the model.

As in §4, we give a comparison of this model with other models where we included/excluded the trend, and modelled the errors $\{\varepsilon_t\}_t$ using: (i) iid converse $GEV(\gamma, \mu, \sigma)$; (ii) iid Gaussian; (iii) an AR(1) process $e_t = \phi e_{t-1} + \eta_t$ where $\{\eta_t\}$ iid Gaussian; and (iv) an AR(1) process $e_t = \phi e_{t-1} + \eta_t$ with $\{\eta_t\}$ iid converse $GEV(\gamma, \mu, \sigma)$. We summarize the results in table 3, where the linear trend is excluded and in table 4, where the linear trend is included. The most suitable model, when ozone is included, was found to be

$$y_t = 0.0091t + 6.61 \sin\left(\frac{2\pi}{12}t\right) + 7.22 \cos\left(\frac{2\pi}{12}t\right) - 0.0267x'_{t-1} + e_t, \quad (5.2)$$

$$e_t = 0.562e_{t-1} + \eta_t, \quad \text{where } \eta_t \sim \text{converse } GEV(-0.0969, -5.67, 3.59).$$

In figure 5, we plot the one-step ahead forecast of minimum temperatures using the previous ozone values. The estimated mean squared error of the forecasts calculated using model (4.2) (where the ozone is excluded) is $\hat{\sigma}^2 = 11.09$ (see table 2), whereas using model (5.1), where the ozone is included, it is 10.14 (see table 4). Therefore, the forecasts obtained for the minimum temperatures, using ozone as an auxiliary variable, increase the quality of the forecasts. Furthermore, we note that there is a small decrease in the magnitude of the

Table 3. The temperatures from 1958 to 2003 modelled using seasonal and ozone terms.

parameter/s.e.	ζ	A	B	ϕ	α	μ	σ	γ	BIC	prediction MSE
independent GEV		5.51 (0.395)	5.50 (0.309)		-0.014 (0.0097)	-10.8 (0.231)	4.49 (0.172)	0.005 (0.040)	6.16	13.47
independent Gaussian	-13.4 (0.236)	7.60 (0.334)	7.18 (0.334)		-0.344 (0.0131)				6.25	15.60
dependent GEV		6.327 (0.493)	7.463 (0.405)	0.614 (0.030)	-0.0163 (0.0076)	-3.76 (0.445)	3.62 (0.124)	-0.085 (0.0294)	5.63	11.59
dependent Gaussian	-13.4 (0.534)	7.575 (0.462)	7.137 (0.461)	0.674 (0.0314)	-0.132 (0.0086)				5.69	10.69

Table 4. The temperatures from 1958 to 2003 modelled using a seasonal term, a linear trend and an ozone term.

parameter/s.e.	ζ	A	B	β	ϕ	α	μ	σ	γ	BIC	prediction MSE
independent GEV		6.326 (0.364)	5.82 (0.305)	-0.0129 (0.0013)		-0.223 (0.0099)	-15.9 (0.364)	4.24 (0.305)	-0.053 (0.0099)	5.98	19.6
independent Gaussian	-19.04 (0.517)	7.65 (0.298)	7.16 (0.298)	0.0156 (0.0013)		-0.0326 (0.0117)				6.02	24.6
dependent GEV		6.61 (0.4630)	7.22 (0.382)	0.0091 (0.0023)	0.562 (0.0307)	-0.0267 (0.0074)	-5.67 (0.631)	3.59 (0.121)	-0.0969 (0.027)	5.60	10.14
dependent Gaussian	-19.0 (1.01)	7.61 (0.422)	7.16 (0.422)	0.015 (0.0026)	0.591 (0.0344)	-0.0145 (0.0087)				5.64	11.17

coefficient (β) of the linear trend, and this may be due to the ozone variable. This may indicate that the ozone levels can be an influencing factor for global warming in the Antarctic region.

6. Discussion and summary

We see from the models fitted (see tables 1–4) that the minimum/maximum temperatures can be explained by linear time-series models with innovations having extreme value distributions. This is confirmed by the probabilistic plots. We observe an upward trend in the minimum temperatures (approx. 6.7°C over 53 years), and no significant increase in the maximum temperatures. If this is confirmed, this may have serious effect on climate change in the Antarctic region and may affect the rest of marine life and ecosystems (see Smith *et al.* 2003). Our conclusions drawn on the basis of our model applied to minimum/maximum temperatures seem to be similar to those drawn by Turner *et al.* (2005) on the basis of average monthly temperatures; but our analysis identifies this increase to be due to an increase in minimum temperatures and also relates this to the ozone levels. The forecasts we obtained from January to December 2004 are very close to the true values observed, which seems to confirm that our models are appropriate. Indeed, the forecasting performance for minimum temperature improved when we included in the model the ozone levels. This, together with the cross-correlation analysis, indicates that the ozone levels may have some effect on the minimum temperatures. The minimum/maximum data are not sufficiently long to draw any long-term valid conclusions, but we believe that, in the future, these methods can be used by climatologists to obtain better forecasts when more observations are available.

Several interesting modelling questions arise from our analysis. The first is how one should model the dependence structure of extreme value observations. Obviously, using an ARMA process with (conv) GEV innovations may not be the only way to do this. The second is an open question whether a random variable, which has an ARMA representation with (conv) GEV innovations, also belongs to the (conv) GEV family of distributions.

We wish to thank Dr Richard Chandler, University College, London, for many useful suggestions. We wish to thank very sincerely Drs Steve Colwell, Jon Shanklin and John Turner of the British Antarctic Survey, Cambridge, for providing us the data and always willingly helping with questions posed by one of the authors (T.S.). We would like to thank the EPSRC for supporting the PhD studies of G.L.H. Finally, we would like to thank two anonymous referees for giving several useful suggestions which helped to improve the presentation of the paper.

Appendix A. Extreme value distributions

In this appendix, we give a brief description of the properties of extreme value distributions used in this paper. Suppose we have a sequence of independent identically distributed random variables with cumulative distribution function $F(x)$. Let $M_n = \max(X_1, X_2, \dots, X_n)$. Then it is well known that

$$\Pr\{M_n \leq x\} = F^n(x),$$

and for constants $a_n > 0$ and b_n , so that as $n \rightarrow \infty$

$$\Pr\left\{\frac{(M_n - b_n)}{a_n} \leq x\right\} = F^n(a_n x + b_n) \rightarrow H(x). \quad (\text{A } 1)$$

When $H(x)$ is of the form

$$H(x) = \exp\left\{-\left[1 + \gamma\left(\frac{x - \mu_X}{\sigma}\right)\right]^{-1/\gamma}\right\}, \quad 1 + \gamma\left(\frac{x - \mu_X}{\sigma}\right) > 0,$$

we say the distribution is a generalized extreme value (GEV) distribution. Here, μ is a location parameter, $\sigma > 0$ is a scale parameter and γ is the shape parameter. The above distribution is often used in modelling data from various disciplines (e.g. [Smith 2001](#)). The three special cases of $H(x)$ are Gumbel, Frechet and Weibull (cf. [Gumbel 1958](#)), which are obtained by choosing the parameter γ appropriately.

The expected value and variance of the random variable X with the above distribution are given by ([Laycock et al. 1990](#))

$$\begin{aligned} E(X) &= \begin{cases} \xi + \frac{\sigma}{\gamma} \Gamma(1 - \gamma); \gamma < 1 \\ \mu + \varrho\sigma; \gamma = 0 \end{cases} \\ \text{Var}(X) &= \begin{cases} \frac{\sigma^2}{\gamma^2} [\Gamma(1 - 2\gamma) - \Gamma^2(1 - \gamma)]; \gamma < \frac{1}{2} \\ \frac{\sigma^2 \pi^2}{6}; \gamma = 0 \end{cases} \end{aligned} \quad (\text{A } 2)$$

where $\xi = \mu - (\sigma/\gamma)$ and ϱ is the Euler's constant.

We are often interested in the limiting probability distribution of $\min(x_1, x_2, \dots, x_n) = -\max(-x_1, -x_2, \dots, -x_n)$, which is given by

$$\begin{aligned} F_Y(y) &= 1 - H(-x) \\ &= 1 - \exp\left\{-\left[1 + \gamma\left(\frac{-y + \mu_Y}{\sigma}\right)\right]^{-1/\gamma}\right\}; \quad 1 + \gamma\left(\frac{-y + \mu_Y}{\sigma}\right) > 0, \end{aligned}$$

where $\mu_Y = -\mu_X$ and the corresponding probability density function is given by

$$f_\eta(y) = \frac{1}{\sigma} \exp\left\{-\left[1 + \gamma\left(\frac{-y + \mu_Y}{\sigma}\right)\right]^{-1/\gamma}\right\} \left[1 + \gamma\left(\frac{-y + \mu_Y}{\sigma}\right)\right]^{-(1+(1/\gamma))}; \quad (\text{A } 3)$$

$$\text{for } 1 + \gamma\left(\frac{-y + \mu_Y}{\sigma}\right) > 0.$$

The distribution (A 3) is known as the converse GEV distribution, conv GEV (γ, μ, σ) . Suppose the random variable $Y \sim \text{conv GEV}(\gamma, \mu, \sigma)$, then the mean of Y is given by $E(Y) = \mu + (\sigma/\gamma)\{1 - \Gamma(1 - \gamma)\}$ and the variance of Y is given by (A 2).

References

- Brockwell, P. J. & Davies, R. A. 1996 *Time series: theory and methods*. New York, NY: Springer.
- Budd, W. F., Jenson, D., Mavrakakis, E. & Coutts, B. 1994 Modelling the Antarctic ice-sheet through time. *Ann. Glaciol.* **20**, 291–297.
- Chatfield, C. 2004 *The analysis of time series. An introduction*. London, UK: Chapman & Hall.
- D’agostino, R. B. & Stephens, M. A. 1986 *Goodness-of-fit techniques*. New York, NY: Marcel Dekker, Inc.
- Davison, A. C. 2003 *Statistical models*. Cambridge, UK: Cambridge University Press.
- Gumbel, E. J. 1958 *Statistics of extremes*. New York, NY: Columbia University Press.
- Hannan, E. J. & Rissanen, J. 1982 Recursive estimation of mixed autoregressive-moving average order. *Biometrika* **69**, 81–94. (doi:10.2307/2335856)
- Houghton, J., Ding, Y., Griggs, D. & Noguer, M. 2001 *Climate change*. Scientific Basis, IPCC Report Cambridge, UK: Cambridge University Press.
- Hughes, G. L. 2002 Multiple regression with correlated, non Gaussian errors—an analysis of extreme temperature data. Ph.D. thesis. University of Manchester Institute of Science and Technology, Manchester.
- Hulbe, C. L. 1997 Recent changes to Antarctic Peninsula ice shelves: what lessons have been learned? *Nat. Sci.* **1**, article 6.
- Huybrechts, P. & Oerlemans, J. 1990 Response of the Antarctic ice sheet to future greenhouse warming. *Clim. Dyn.* **5**, 93–102.
- Kavalieris, L. & Hannan, E. J. 1994 Determining the number of terms in a trigonometric regression. *J. Tm. Ser. An.* **15**, 613–625.
- King, J. C. 1994 Recent climatic variability in the vicinity of the Antarctic Peninsula. *Int. J. Climatol.* **14**, 357–369.
- King, J. C. & Harangozo, S. A. 1998 Climate change in the Western Antarctic Peninsula since 1945: observations and possible causes. *Ann. Glaciol.* **27**, 571–575.
- Laycock, P. J., Cottis, R. A. & Scarf, P. A. 1990 Extrapolation of extreme pit depths in space and time. *J. Electrochem. Soc.* **137**, 64–69. (doi:10.1149/1.2086440)
- Quinn, B. G. & Fernandes, J. M. 1991 A fast and efficient technique for the estimation of frequency. *Biometrika* **78**, 489–498. (doi:10.2307/2337018)
- Smith, R. L. 1985 Maximum likelihood estimation in a class of nonregular cases. *Biometrika* **72**, 67–91. (doi:10.2307/2336336)
- Smith, R. L. 2001 Extreme value statistics in meteorology and environment part 1. Notes from CBMS course in environmental statistics, University of Washington.
- Smith, R., Fraser, W. R., Stammerjohn, S. E. & Vernet, M. 2003 Palmer long-term ecological research on the antarctic marine ecosystem. *Antarct. Res. Ser.* **79**.
- Turner, J., King, J. C., Lachlan-Cope, T. A. & Jones, P. D. 2002 Recent temperature trends in the Antarctic. *Nature* **418**, 291–292. (doi:10.1038/418291b)
- Turner, J., Colwell, S. R., Marshall, G. J., Lachlan-Cope, T. A., Carleton, A. M., Jones, P. D., Lagin, V., Reid, P. A. & Igovkina, S. 2004 The SCAR Reader Project: towards a high quality database of mean Antarctic meteorological observations. *J. Clim.* **17**, 2890–2898. (doi:10.1175/1520-0442(2004)017<2890:TSRPTA>2.0.CO;2)
- Turner, J., Colwell, S. R., Marshall, G. J., Lachlan-Cope, T. A., Carleton, A. M., Jones, P. D., Lagin, V., Reid, P. A. & Igovkina, S. 2005 Antarctic climate change during the past 50 years. *J. Climatol.* **25**, 279–294. (doi:10.1002/joc.1130)