

Extreme events in total ozone over Arosa – Part 1: Application of extreme value theory

H. E. Rieder¹, J. Staehelin¹, J. A. Maeder¹, T. Peter¹, M. Ribatet², A. C. Davison², R. Stübi³, P. Weihs⁴, and F. Holawe⁵

¹Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

²Institute of Mathematics, EPF Lausanne, Lausanne, Switzerland

³Federal Institute of Meteorology and Climatology (MeteoSwiss), Payerne, Switzerland

⁴Institute for Meteorology, University of Natural Resources and Applied Life Sciences (BOKU), Vienna, Austria

⁵Institute for Geography and Regional Research, University of Vienna, Vienna, Austria

Received: 22 April 2010 – Published in Atmos. Chem. Phys. Discuss.: 18 May 2010

Revised: 31 August 2010 – Accepted: 22 September 2010 – Published: 25 October 2010

Abstract. In this study ideas from extreme value theory are for the first time applied in the field of stratospheric ozone research, because statistical analysis showed that previously used concepts assuming a Gaussian distribution (e.g. fixed deviations from mean values) of total ozone data do not adequately address the structure of the extremes. We show that statistical extreme value methods are appropriate to identify ozone extremes and to describe the tails of the Arosa (Switzerland) total ozone time series. In order to accommodate the seasonal cycle in total ozone, a daily moving threshold was determined and used, with tools from extreme value theory, to analyse the frequency of days with extreme low (termed ELOs) and high (termed EHOs) total ozone at Arosa. The analysis shows that the Generalized Pareto Distribution (GPD) provides an appropriate model for the frequency distribution of total ozone above or below a mathematically well-defined threshold, thus providing a statistical description of ELOs and EHOs. The results show an increase in ELOs and a decrease in EHOs during the last decades. The fitted model represents the tails of the total ozone data set with high accuracy over the entire range (including absolute monthly minima and maxima), and enables a precise computation of the frequency distribution of ozone mini-holes (using constant thresholds). Analyzing the tails instead of a small fraction of days below constant thresholds provides deeper insight into the time series properties. Fingerprints of dynamical (e.g. ENSO, NAO) and chemical features (e.g. strong polar vortex ozone loss), and major vol-

canic eruptions, can be identified in the observed frequency of extreme events throughout the time series. Overall the new approach to analysis of extremes provides more information on time series properties and variability than previous approaches that use only monthly averages and/or mini-holes and mini-highs.

1 Introduction

Previous studies have examined the transient low ozone events termed ozone mini-holes (e.g. James, 1998; Krzyścin, 2002; Bojkov and Balis, 2001; Brönnimann and Hood, 2003; Koch et al., 2005; Anton et al., 2008). Two mechanisms have been proposed as important for their formation in the mid-latitudes: (a) far-range meridional transport of air masses from regions with different climatological ozone mixing ratios (e.g. Koch et al., 2005) and (b) local adiabatic vertical displacement of isentropes (Wohltmann et al., 2007). Iwao and Hirooka (2006) showed that ozone mini-holes in the Northern Hemisphere are associated with a combination of two phenomena: pole-ward transport in the lowermost stratosphere and equator-ward transport in the middle stratosphere; and vertical uplift in the lower and middle stratosphere.

The frequency of occurrence of mini-holes and mini-highs (see Koch et al., 2005) depends strongly on their definition, which we now briefly review. James (1998) used a deviation of -70 DU from ozone mean values as a threshold definition of mini-holes and included seasonal variability in the analysis. Bojkov and Balis (2001) applied a fixed threshold for ozone mini-holes of 220 DU over areas larger than 500 000 km². Krzyścin (2002) analyzed long-term changes



Correspondence to: H. E. Rieder
(harald.rieder@env.ethz.ch)

in ozone mini-hole frequency over the Northern Hemisphere, using a threshold for low total ozone defined as a deviation of minus 20 percent from the climatological mean value. Brönnimann and Hood (2003) analyzed the frequency of low-ozone events over north-western Europe in 1952–1963 and 1990–2000. They used a threshold for days with very low total ozone (<225 DU) and found that the frequency of low-ozone events over north-western Europe was quite similar for autumns in 1952–1963 and 1990–2000. For the more recent period they found a significant increase in days with low total ozone during winter, and reported that changes in atmospheric circulation play a major role for the formation of low ozone events. Changes in mean ozone column due to human activity also play an important role, but their extent remains uncertain. Koch et al. (2005) applied a threshold of plus (mini-high) or minus (mini-hole) one standard deviation from the monthly mean in their study. This method has been updated by Antón et al. (2008), applying a more restrictive definition by using a two standard deviation threshold.

Here we show by statistical analysis that the mentioned above concepts assuming a Gaussian distribution of total ozone (e.g. fixed deviations from mean values) do not adequately address the structure of the extremes (see Sect. 2.2). Therefore, in this study we use ideas from extreme value theory (e.g. Coles, 2001; Ribatet, 2007) to compute thresholds for extreme low (termed ELOs) and extreme high (termed EHOs) total ozone. A heavy-tail focused approach is used entailing the fitting of the Generalized Pareto Distribution (GPD) to the Arosa time series. This provides a mathematically well-defined approach to extremes. Further we analyse the frequency distribution of ELOs and EHOs in the context of atmospheric processes.

2 Measurements and methods

2.1 The Arosa long-term ozone time series

Long term measurements of total ozone extending back before 1958 are available only from a few stations in the Northern Hemisphere (see Vogler et al., 2007). The world's longest available total ozone record is from Arosa, Switzerland ($46^{\circ}46' \text{ N}$, $9^{\circ}40' \text{ E}$), where total ozone measurements have been made since late 1926. Total ozone measurements at Arosa were started in 1926 by F. W. P. Götze, continued by G. Perl and then by H. U. Dütsch, and since 1988 have been performed by the Swiss Federal Institute of Meteorology and Climatology (MeteoSwiss). The total ozone record from Arosa is based on direct sun measurements with Dobson spectrophotometers and has been homogenized (Staelin et al., 1998a). Total ozone data from Arosa used in this study was provided by the World Ozone and Ultraviolet Radiation Data Centre (WOUDC, <http://www.woudc.org>). In this study we use daily mean values, available on approximately 250 days per year (Staelin, 1998a).

Annual averages of total ozone from 1927–2008 and the 11-year running mean are shown in Fig. 1. While no trend can be found in the earlier parts of the record (Lund et al., 1995) a negative trend in total ozone over Arosa is clearly visible from the mid 1970s until the early 1990s (e.g. Staelin et al., 1998b).

2.2 Frequency distribution of total ozone at Arosa

In the previous studies mentioned above (see Sect. 1), thresholds for ozone mini-holes and mini-highs have been based on fixed values, on fixed deviations from monthly/seasonally/yearly mean values, or on 1σ or 2σ deviations from these mean values. From a statistical point of view this is appropriate if the data series can be approximated by a Gaussian distribution. To test total ozone data for Gaussian behaviour a simple standard test, the Shapiro-Wilk Test, was performed for the periods 1927–1959 (pre-industrial conditions with respect to stratospheric chlorine and volcanic disturbance) and 1960–2008 (industrial conditions with respect to stratospheric chlorine and volcanic disturbance). The test values (p -value, a statistical measure of evidence against the null hypothesis, following, Royston, 1982a, b) of 10^{-5} (1927–1959) and 10^{-15} (1960–2008) show clearly that daily values of total ozone are not adequately described by a Gaussian distribution. This test was also applied to deseasonalized and detrended data, obtained by applying STL: Seasonal Trend decomposition of the time series based on LOESS (LOcally wEighted Scatterplot Smoothing) (e.g. Cleveland et al., 1990); for details see Sect. 3.1. Here the p -values of 10^{-6} (1927–1959) and 5×10^{-8} (1960–2008) show clearly that detrended and deseasonalized daily values of total ozone are not adequately described by a Gaussian distribution.

In Fig. 2 the deviation of total ozone ((a–b) original observations and (c–d) detrended and deseasonalized observations) from a Gaussian distribution is shown in Quantile-Quantile-plots (QQ-plots) (which compare two probability distributions) for the two periods described above. Data from a Gaussian distribution would lie close to the diagonal line, but for both periods total ozone shows strong positive skewness, evidenced by the departure from linearity, and hence are clearly non-Gaussian, as was shown by the very significant Shapiro-Wilk statistic. Several studies (e.g. Davison and Smith, 1990; Coles, 2001; Coles and Pericchi, 2003) showed that proper analysis of the tails of such datasets requires extreme value theory, and here we explore its use to describe extreme events in total ozone and to select thresholds based on the tails of the ozone data.

2.3 The Generalized Pareto Distribution (GPD)

The Generalized Pareto Distribution (GPD), one of the most commonly used distributions in the framework of extreme value theory (e.g. Davison and Smith, 1990; Cooley et al.,

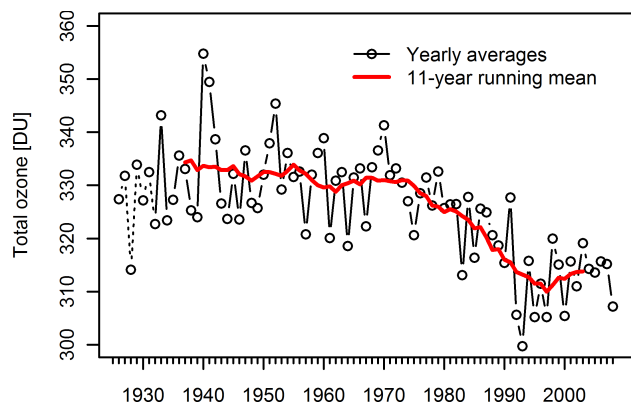


Fig. 1. Annual averages and 11-year running mean of total ozone over Arosa, Switzerland 1927–2008. Note: dotted line for annual averages marks period with large fractions of missing data before 1932.

2007; Ribatet et al., 2009), is named after the Italian mathematician and economist Vilfredo Pareto. Although originally derived in a different context, it also arises as the natural distribution for the exceedance of a random variable over a high threshold, and hence is widely used in the modeling of rare events (e.g. Davison and Smith, 1990). Extreme value theory has been applied to analyze rare events in various fields of climate science (e.g. Fawcett and Walshaw, 2006; Cooley et al., 2007; Ribatet et al., 2009). In this study extreme value theory is used for the first time in the field of total ozone research.

Here we describe the theoretical arguments for modelling exceedances over a threshold. However, the modelling procedure for values (x) below a low (enough) threshold is the same (see Eq. 1); to apply to low values: negate the values, apply methods for modelling beyond a high threshold, and then negate the results, because

$$\min(x_i) = -\max(-x_i). \quad (1)$$

Arguments on the asymptotic behaviour of the data series (Pickands, 1975) justify the use of the GPD for modelling exceedances over a sufficiently high threshold (Coles, 2001). More precisely, the GPD is the limiting distribution of normalized exceedance over a threshold, as the threshold approaches the maximum of the distribution. The GPD is defined as:

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

$$\sigma > 0, x > u, 1 + \xi \frac{x-u}{\sigma} > 0.$$

where u is the threshold, and σ and ξ are the scale and shape parameters. The threshold u determines which levels are to be considered as extreme. The scale parameter σ is a measure of the spread of the distribution, with the same dimensions as the variable x ; the larger σ , the more spread out the

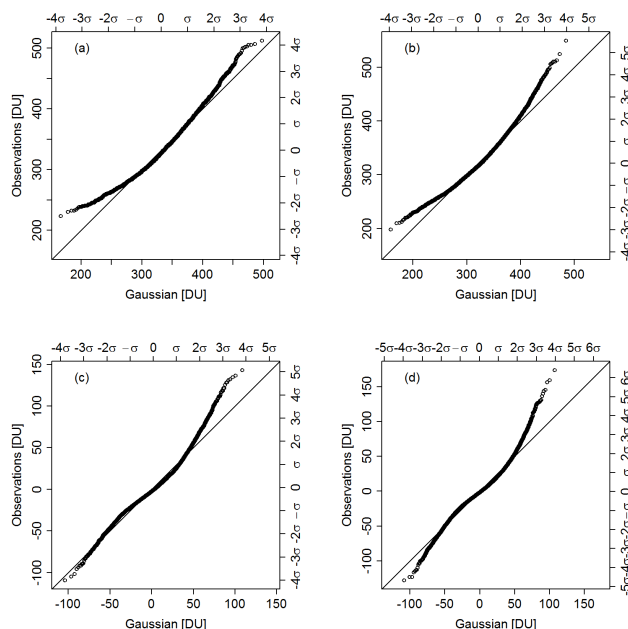


Fig. 2. Quantile-Quantile plots comparing (a) total ozone observations at Arosa (Switzerland) with a Gaussian distribution (least-mean-square fitted to observations) for the period 1927–1959, (b) total ozone observations at Arosa (Switzerland) with a Gaussian distribution (least-mean-square fitted to observations) for the period 1960–2008, (c) remainder (deseasonalized and detrended total ozone observations (by LOESS (STL) see Sect. 3.1 and Fig. 3d) at Arosa (Switzerland)) with a Gaussian distribution (least-mean-square fitted to observations) for the period 1927–1959 and (d) remainder (deseasonalized and detrended total ozone observations (by LOESS (STL) see Sect. 3.1 and Fig. 3d) at Arosa (Switzerland)) with a Gaussian distribution (least-mean-square fitted to observations) for the period 1960–2008.

distribution is. The shape parameter ξ determines the shape of the tail of the distribution rather than simply shifting it (as u does) or stretching/shrinking it (as σ does). Selection of the threshold value u is described in Sect. 2.4 and Sect. 3.2 below.

2.4 Threshold selection

In practice, the GPD parameters are fitted to observed positive exceedances $x - u$ by maximum likelihood estimation, though other estimators, such as probability weighted moments estimators, are sometimes used.

An essential preliminary step is to determine an appropriate threshold u for which the asymptotic GPD approximation holds. This requires consideration of the trade-off between bias and variance: too high a threshold will reduce the number of exceedances and thus increase the estimation variance (see Fig. 6, discussed later), whereas too low a threshold will induce a bias because the GPD will fit the exceedances poorly (see Fig. 7, discussed later).

Daily total ozone data show serial correlation and therefore autocorrelation has to be taken into account in the analysis. Fawcett (2005) showed that if one ignores temporal dependence within a data set, estimates for the GPD parameters are still consistent but the standard errors have to be inflated. However, to work in a more conventional statistical framework (i.e. independent realizations) we decided to take autocorrelation into account by means of a declustering procedure, which was applied prior to threshold selection. The method applied (as suggested by Davison and Smith, 1990) was to formally identify clusters of extremes and then use only the maximum (minimum) of each cluster in the analysis. A cluster is defined as consecutive exceedances above/below a certain threshold u . Once an observation falls below/above u the cluster is deemed to have terminated. A declustering parameter κ (a separation interval between neighboring clusters) is introduced to mitigate the potential dependence between clusters. For the analysis $\kappa = 1$ day was sufficient to remove any serial dependence.

Tools for threshold choice (e.g. Coles, 2001; Ribatet, 2007) involve comparing the theoretical behaviour of the GPD and the empirical behaviour of the observations. A first tool, the Mean Residual Life (MRL) plot, entails plotting the mean of the exceedances over the threshold, as the threshold varies. If the observations follow a GPD with $\xi < 1$, the mean exceedance should vary linearly with the threshold. More precisely, we graph the points as defined below and look for the smallest threshold u above which the graph is roughly linear:

$$\left\{ \left(u, \frac{1}{n_u} \sum_{i=1}^{n_u} [x_i - u] \right), u < \max(x_i) \right\},$$

where n_u is the number of exceedances above u and x_i is the i -th observation above u . A second tool, the Threshold Choice plot (TC-plot), uses the fact that if the exceedances above a threshold u are GPD with scale $\sigma(u)$ and shape $\xi(u)$, then exceedances above a larger threshold u' are also GPD but with scale $\sigma(u) + \xi(u)(u' - u)$ and shape $\xi(u)$. Using this property, we plot the points defined by

$$\{(u, \sigma^*(u)), u < \max(x_i)\},$$

$$\{(u, \xi(u)), u < \max(x_i)\},$$

and look for the smallest threshold u above which the graphs are roughly constant:
where

$$\sigma^*(u) = \sigma(u) - \xi(u)u \quad (3)$$

is called the modified scale parameter. In our application to the ozone time series, the threshold values for extreme low and extreme high total ozone were calculated with the above tools using the package POT (Ribatet, 2007) in the statistical computing environment R (R Development Core Team, 2008).

Starting from the 1970s, anthropogenic emission of ozone depleting substances (ODS) had a major influence on the Earth's ozone layer, and one might argue that appropriate threshold values for natural extremes in total ozone must be based on measurements prior to 1970. Moreover, total ozone is also influenced by changes in quantities such as the aerosol loading of the Earth's atmosphere. The Northern Hemisphere aerosol climatology (e.g. Sato et al., 1993) shows some smaller peaks in the 1960s. From these two “constraints” total ozone data at Arosa is considered as anthropogenically and volcanically unperturbed before the 1960s.

As the aim of this study is to address the influence of short-term and long-term processes (including long-term trends) on the variation of extremes, analyzing an anthropogenically and volcanically unperturbed period (1927–1959) and an anthropogenically and volcanically perturbed period (1960–2008) is from the authors point of view a solid way to address ozone changes within two clearly different regimes.

We do not detrend the data or use different sub-periods for threshold selection because this would require a subjective a priori decision on the type of trend and/or time periods. The influence of such decisions on the results is hard to quantify. More extensive discussion on the detrending of the data and the choice of different sub periods is presented in the supplementary material of this paper. Figure S1 in the supplementary material provides an example on how strongly a trend estimator is influenced by the time window chosen. Therefore we prefer to consider the observed trend within our analysis and not within a preceding step. Additionally Fig. S2 in the supplementary material compares the thresholds used in this study with thresholds chosen on other (shorter) time periods. Figures S3 and S4 in the supplementary material compare the influence of thresholds selected on different time periods on the frequency distribution of ELOs and EHOs. These figures show that the differences in the frequency of EHOs and ELOs are quite small and that the main findings of the study are robust and do not depend on the time windows used for threshold determination. As consequence we prefer to use the longest (unperturbed and perturbed) time windows (1927–1959 and 1960–2008), to compute thresholds for extreme low and high total ozone, for statistical reasons.

3 Results and discussion

3.1 Seasonal decomposition of the Arosa total ozone time series

The information contained in the Arosa total ozone record may be exploited in more detail in an STL-plot (e.g. Cleveland et al., 1990), the Seasonal Trend decomposition of the time series based on LOESS (LOcally wEighted Scatterplot Smoothing (e.g. Cleveland, 1979; Cleveland and Devlin, 1988; Cleveland et al., 1990). This is one of the many and now widely-used “modern” statistical modelling methods

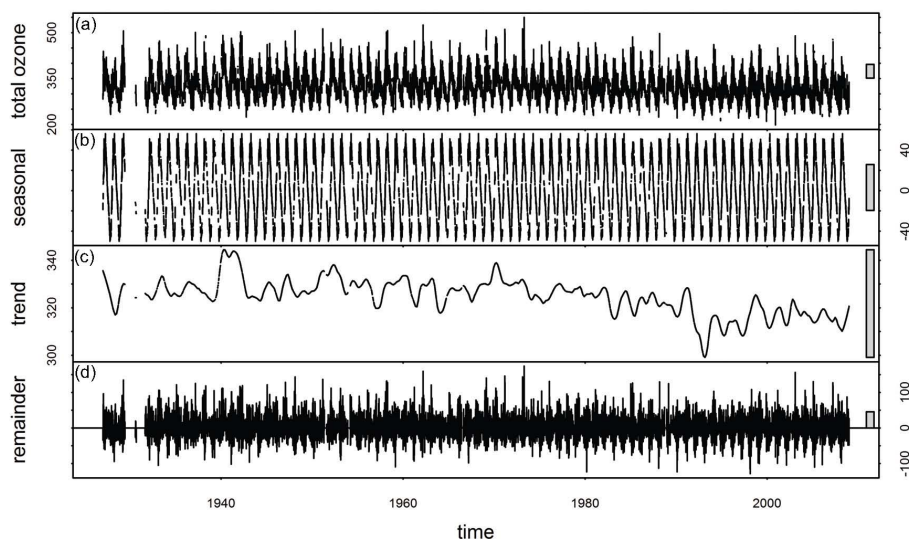


Fig. 3. Decomposition plot of the Arosa total ozone record based on a Seasonal-Trend decomposition procedure based on LOESS (STL): (a) daily mean values, (b) seasonal component, (c) trend component, (d) remainder. For convenient reference, right-hand grey bars in (a–d) refer to the same magnitude (40 DU). Gaps in the time series indicate periods with missing data. All panels are for 1927–2008. LOESS was computed using the `stl-R` function with parameters `t.window=549`, `t.degree=1`, `s.window="periodic"`, and `s.degree=0`.

that build on “classical” methods such as linear and nonlinear least squares regression (Fig. 3). All daily mean values of the homogenized total ozone record are plotted in Fig. 3a, which shows that there is large variability within the individual daily mean total ozone measurements. In Fig. 3b the seasonal component of the Arosa time series is extracted; there is a clear seasonal cycle in total ozone, with maxima in spring and minima in late fall/early winter (in accordance with the Brewer-Dobson circulation transporting ozone-rich air towards northern mid-latitudes during boreal winter). Fig. 3c shows the LOESS-smoothed trend component for the Arosa time series. During the first 45 years there is no clear trend, but from the 1970s onwards total ozone decreases, reaching its lowest values in the early 1990s after the volcanic eruption of Mount Pinatubo. After a period of increase after 1992, the total ozone values remained subsequently rather stable at a low level. The remainder of the Arosa time series, i.e. the deviation from trend and estimated seasonal component, plotted in Fig. 3d, shows that the estimation of the seasonal cycle is not perfect, because the remainder shows weak seasonality. It also shows a high degree of short-term variability. Inspection of the mean adjusted residuals shows that their distribution changed from a more asymmetric shape in the earlier years to a more symmetric shape in the later periods (see Fig. 3d). Thus the distribution of (deseasonalized and detrended) total ozone has changed with time, a feature not represented in standard monthly mean analyses.

The grey bars at the right of all four subpanels of Fig. 3 aid in comparing the relative sizes of the trend, seasonal, and remainder components of the decomposition. Figure 3a and c show that the entire (negative) trend in total ozone over Arosa is small compared to the natural variability in the time series.

3.2 Thresholds for low and high total ozone

Following Sect. 2.4, Fig. 4a–f provides examples (for March) of the Mean Residual Life plots (MRL-plot) and the Threshold Choice plots (TC-plots). In the MRL-plots (Fig. 4a and d), points defined by

$$\{(u, \sigma^*(u)), u < \max(x_i)\},$$

are plotted for the ozone data set. As stated above, if the observations follow a GPD with $\xi < 1$, the mean exceedance should decrease linearly with the threshold value. In the examples provided in Fig. 4a (for the period 1927–1959) and Fig. 4d (for the period 1960–2008) we see that linear behaviour ends around 400 DU or 390 DU, respectively. Next the information gained from the MRL-plot is analyzed using TC-plots (Fig. 4b–c for the period 1927–1959 and Fig. 4e–f for the period 1960–2008). While Fig. 4b and e show the original shapes, Fig. 4c and f show the modified scale parameter (see Eq. 3) at certain threshold values. In the TC-plot a constant shape can be identified starting at 400 DU (period 1927–1959) and 390 DU (period 1960–2008), corroborating the MRL-results.

Threshold selection requires consideration of the trade-off between bias and variance. Above we identified the

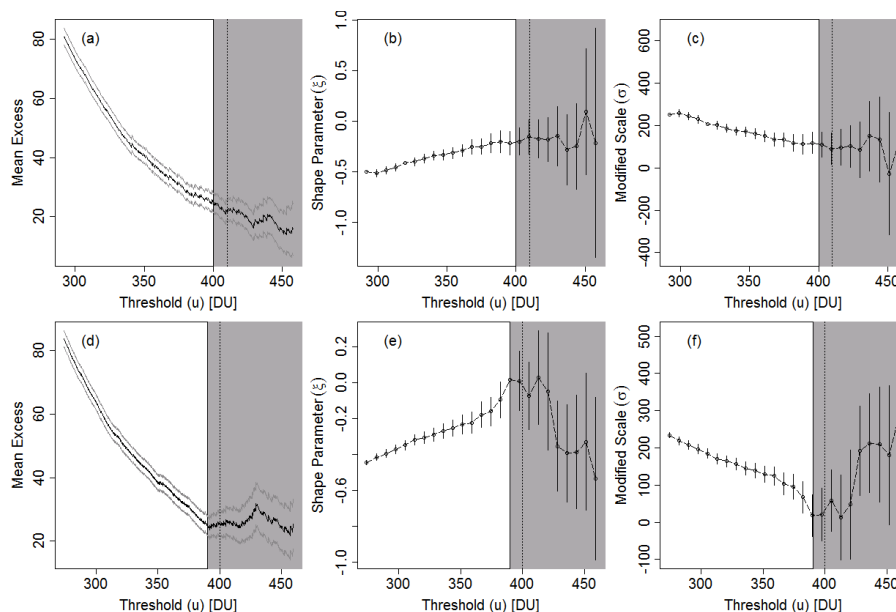


Fig. 4. Threshold selection plots (examples for high total ozone in March 1927–1959): **(a)** Mean Residual Life Plot (MRL-Plot) and **(b–c)** stability plots (TC-plots) computed with POT-package (Ribatet, 2007). **(d–f)** same as (a–c) but for period 1960–2008. Grey area gives stability region and the dashed vertical line gives the selected threshold ($u = 410$ DU period 1927–1959 and $u = 400$ DU period 1960–2008) in all plots.

constant part within the tail with reasonably low bias and variance, namely thresholds from 400 DU (1927–1959) and 390 DU (period 1960–2008) upwards to be considered (see grey boxes). The best choice for the threshold u within these ranges can be selected by the comparison of observed and modeled events at each potential threshold value. This comparison is based on three different plots (see Fig. 5): a density plot, a Quantile-Quantile plot (QQ-plot), a probability plot. The QQ-plot (see Fig. 5b and e) compares observed and modeled events, the probability plot (see Fig. 5c and f) compares the empirical and predicted probability for these events, and the density plot compares a non-parametric estimate of the density with the fitted GPD-model (see Fig. 5a and d). This analysis is done for all potential threshold values within the range selected from Fig. 4. The threshold value for which the QQ-plot, the probability plot and the density plot show the best agreement between the observed and modeled values is taken as the threshold for this month.

In Fig. 5 examples of these plots for the month of March are shown for the most suitable threshold value $u=410$ DU in the period 1927–1959 and $u=400$ DU for the period 1960–2008. From these figures the good agreement between GPD and observations for both periods is obvious.

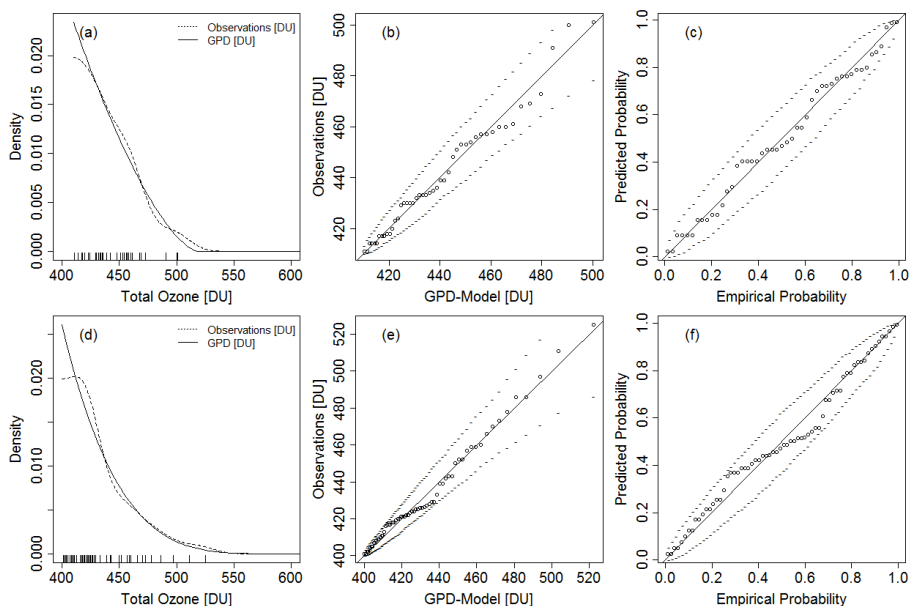
As mentioned above, threshold selection requires consideration of the trade-off between bias and variance. To compare three different models (GPD with different parameters dependent on u , u too high and u too low) on the same basis, we need to use the same set of observations. The methodology used was: (i) to update the GPD parameters

with a threshold equal to 445 DU (443 DU); (ii) to compare the model uncertainties by standard errors (see Table 1) and (iii) assess goodness of fit from QQ-plots and density plots. As expected the standard errors for the models using too high threshold values are much larger than for the optimized model. Although many more observations were used when using a lower threshold, the standard errors from the optimized model might be smaller. This statement combined with QQ-plots indicates that the optimized model corresponds to the best trade-off between bias and variance and has the best predictive capabilities. Additionally Figs. 6 and 7 show the density plots, the QQ-plots and the probability plots for too high and too low threshold values.

The threshold selection was performed (for the periods 1927–1959 and 1960–2008) for low and high total ozone for each calendar month following the procedure described above. The strong seasonal cycle was accommodated by smoothing monthly threshold values by linear interpolation and thus producing daily thresholds. Figure 8 shows these thresholds for extreme low and high total ozone and monthly means for 1927–1959 and 1960–2008. The monthly mean values of the time series (in Figs. 8 and S5) show larger differences than the corresponding thresholds for low and high total ozone. Changes in the frequency of ELOs and EHOs within the perturbed time period cannot be attributed to a particular choice of the reference period, (see supplementary material and “Interactive Discussion” to the discussion paper). Rather, such changes must be interpreted as fingerprints of the physical factors affecting ozone.

Table 1. Scale (σ) and shape (ξ) parameters for models with u selected (410 DU, 400 DU), u too high (445 DU, 443 DU) and u too low (377 DU, 375 DU) for the time periods 1927–1959 and 1960–2008. Standard errors are given in brackets.

period	scale parameter (σ)			shape parameter (ξ)		
	u	u (too high)	u (too low)	u	u (too high)	u (too low)
1927–1959	42.77 (8.13)	27.35 (10.11)	76.26 (11.22)	-0.39 (0.13)	-0.35 (0.29)	-0.58 (0.10)
1960–2008	38.36 (6.45)	46.89 (16.08)	55.87 (7.13)	-0.18 (0.11)	-0.50 (0.26)	-0.31 (0.08)

**Fig. 5.** (a) Tail density plot of observations and GPD for high total ozone ($u = 410$ DU), (b) Quantile-Quantile-comparison between observed and GPD-model predicted extreme values, (c) empirical and predicted probability for these events (all three for the month of March in 1927–1959). (d–f) same as (a–c) but for period 1960–2008 with $u = 400$ DU. In (b–c) and (e–f) dotted lines represent the 95% confidence intervals, straight lines give the identity line.

3.3 Frequency distribution of extreme events in low and high total ozone

In a next step we grouped the total ozone observations over Arosa into three groups using the thresholds selected for 1927–1959 and 1960–2008 (see Eqs. 4–6):

$$\text{ELO} = \{x(t) : x(t) < u_{\text{LOW}}(t)\}, \quad (4)$$

$$\text{EHO} = \{x(t) : x(t) > u_{\text{HIGH}}(t)\}, \quad (5)$$

$$\text{NEO} = \{x(t) : u_{\text{LOW}}(t) \leq x(t) \leq u_{\text{HIGH}}(t)\}. \quad (6)$$

Here, ELO denotes a day with extreme low total ozone, EHO a day with extreme high total ozone, NEO a non-extreme day, $x(t)$ is the total ozone column for a single day, $u_{\text{HIGH}}(t)$ is the threshold for high total ozone for a specific day of the year and $u_{\text{LOW}}(t)$ is the threshold for low total ozone for a specific day of the year.

Figures 9a–d depict a non-parametric estimate of the density of total ozone over Arosa for 1927–1959 and 1960–2008. While Fig. 9a (all data in 1927–1959) and Fig. 9b (example for March in 1927–1959) show estimated densities for the entire dataset, Fig. 9c (all data in 1960–2008) and Fig. 9d (example for March in 1960–2008) show the estimated density of the time series without extremes. We see that total ozone over Arosa shows heavy tails in the data distribution (see Fig. 9a and b) and that the density of the remaining bulk (time series without ELOs and EHOs) is almost equal in both time periods. Figure 9c and d provide a quantitative indication that observed trends in total ozone are strongly dominated by extreme events, as we see that the bulk of the distribution shows no big differences. This issue is discussed in detail in a companion paper (Rieder et al., 2010, hereafter called Part 2).

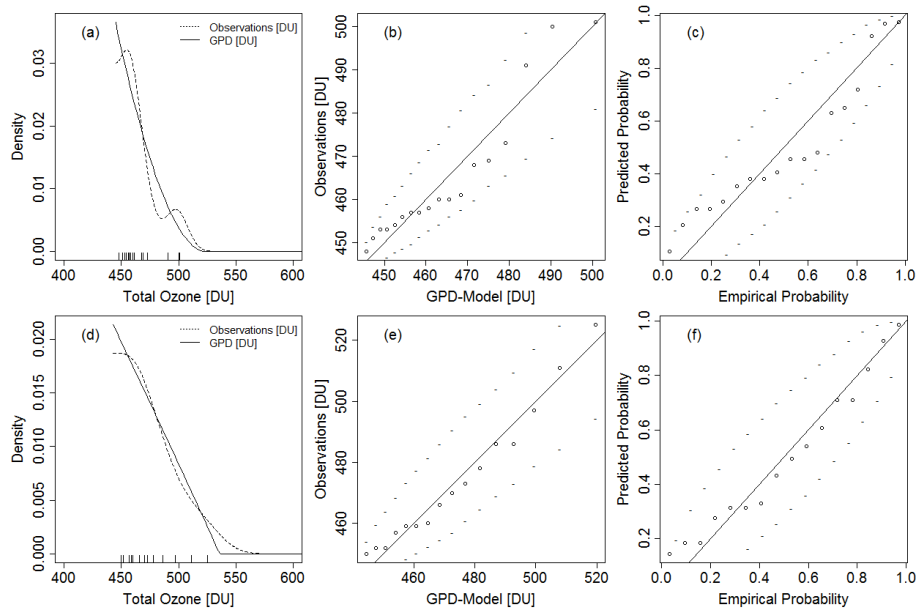


Fig. 6. (a) Tail density plot of observations and GPD for high total ozone with too high threshold values ($u = 445$ DU), (b) Quantile-Quantile-comparison between observed and GPD-model predicted extreme values, (c) empirical and predicted probability for these events (all three for the month of March in 1927–1959). (d–f) same as (a–c) but for period 1960–2008 with $u = 443$ DU. In (b–c) and (e–f) dotted lines represent the 95% confidence intervals, straight lines give the identity line.

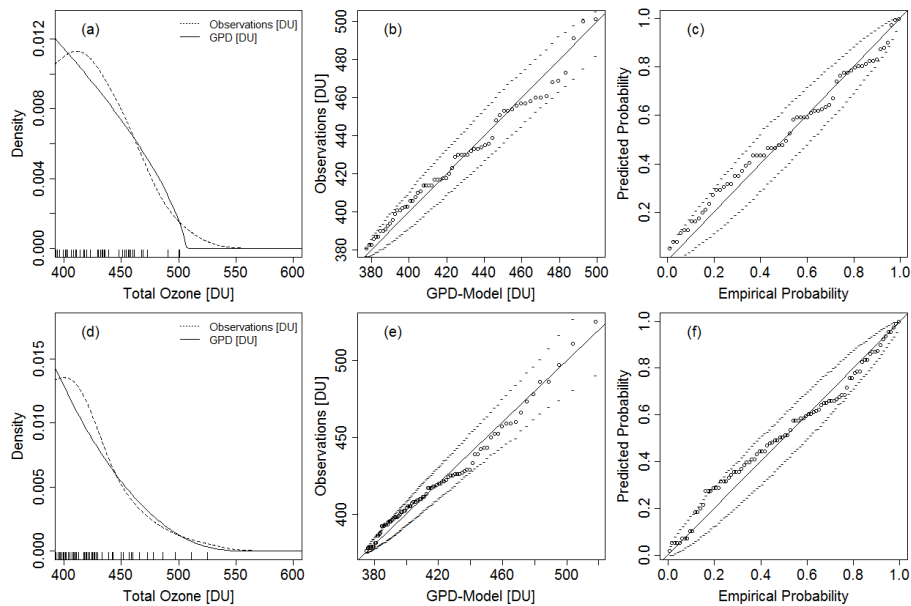


Fig. 7. (a) Tail density plot of observations and GPD for high total ozone with too low threshold values ($u = 377$ DU), (b) Quantile-Quantile-comparison between observed and GPD-model predicted extreme values, (c) empirical and predicted probability for these events (all three for the month of March in 1927–1959). (d–f) same as (a–c) but for period 1960–2008 with $u = 375$ DU. In (b–c) and (e–f) dotted lines represent the 95% confidence intervals, straight lines give the identity line.

The observed frequency for all three groups (EHOs, ELOs and NEOs) is plotted as a time series in Fig. 10 on an annual scale. The first period of the Arosa record reveals some fluctuations in EHOs and ELOs of the total ozone series (ranges

between 5% and 20% in annual means) as well as in the remaining bulk of observations. In 1940–1942 a first strong increase in EHOs appears, linked to a strong ENSO event (for details see Brönnimann et al., 2004; Part 2), followed by

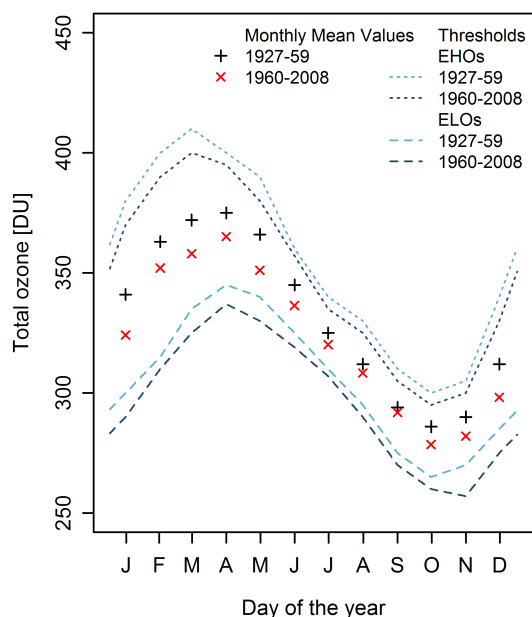


Fig. 8. Thresholds for extreme high (EHOs) and low (ELOs) total ozone and monthly means of total ozone over Arosa, for the periods 1927–1959 and 1960–2008.

a second period of about 30 years with less frequent EHOs. Overall the number of EHOs is slightly larger than that of ELOs during the first 40 years of the Arosa record. From the mid-1970s or early 1980s a decrease in the number of EHOs and an increase in the frequency of ELOs is observed. During the last 5–10 years a decrease in the frequency of ELOs and a stabilization or small increase in the frequency of EHOs can be observed over Arosa. The influence of these extreme events on annual and seasonal mean values and trends in the context of atmospheric chemistry, physics and dynamics is further discussed in a companion paper (see Part 2).

3.4 Application of GPD-model to ozone mini-hole frequency

The frequency distribution of extremes was compared with earlier papers related to ozone mini-holes (e.g. Brönnimann and Hood, 2003; Bojkov and Balis, 2001). Following the definition of Brönnimann and Hood (2003) using a threshold of 225 DU a total number of 15 ozone mini-holes can be found in the Arosa total ozone series (see Table 2). The probability of a single extreme event (ELO) to be counted as an ozone mini-hole was computed from the low heavy tail of the Arosa time series (ELOs). The analysis showed that the probability of observing a mini-hole is only relevant in the period November to February, with highest probability in November and December (not shown here). This agrees well with the number of mini-holes observed at Arosa (compare Table 2). Further the GPD-model was also used to estimate the number of ozone mini-holes following the more

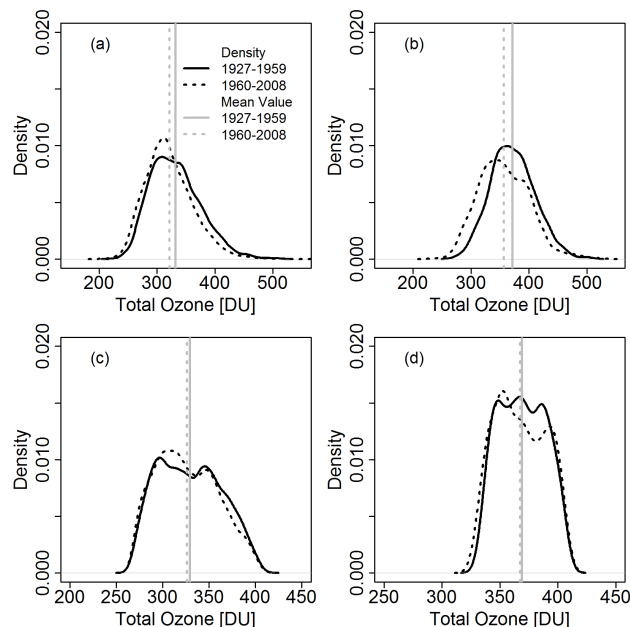


Fig. 9. Density plots of (a) annual total ozone observations at Arosa for the time periods 1927–1959 and 1960–2008, (b) ozone observations for March in the periods 1927–1959 and 1960–2008, (c) annual total ozone observations with extremes removed for the time periods 1927–1959 and 1960–2008 and (d) ozone observations with extremes removed for March in the periods 1927–1959 and 1960–2008. Mean values for the time periods are given as vertical lines.

restrictive definition of Bojkov and Balis (2001) (threshold of 220 DU, see days marked with * in Table 2). The number of predicted mini-holes agrees well with the observations: 16 (for a threshold of 225 DU) and 6 mini-holes (for a threshold of 220 DU) would be expected from the GPD-model, while 15 respectively 6 have been observed. This result shows that the number of mini-holes can be estimated with high accuracy from the GPD-model. For the prediction of the temporal evolution of ozone mini-holes an updated and more complex statistical model would be needed, which is beyond the scope of the present study. Here we show only that the frequency distribution of mini-holes can be estimated highly accurate within the presented framework.

3.5 Atmospheric processes and extreme events in total ozone

Figure 10 contains information related to particular events in the Earth's atmosphere. The coincidence of anomalous frequencies of ELOs/EHOs are related to ENSO anomalies (e.g. Brönnimann et al., 2004; Sassi et al., 2004), North Atlantic Oscillation anomalies (NAO) (e.g. Appenzeller et al., 2000), large volcanic eruptions and strong polar vortex ozone depletion (Wohltmann et al., 2007). The analysis and discussion of the influence of these atmospheric processes is the subject of an accompanying paper (see Part 2).

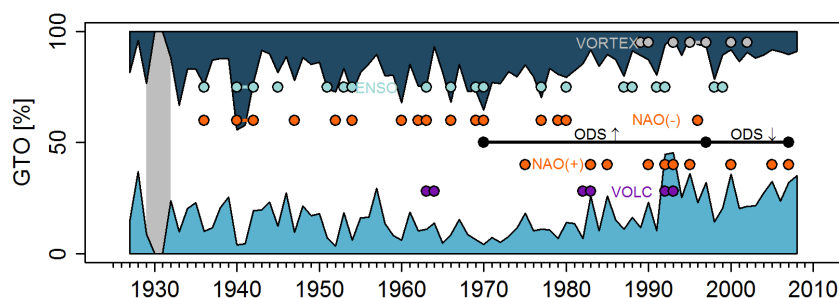


Fig. 10. Annual evolution of fractions of extreme events in low (ELOs – light blue area) and high (EHOs – dark blue area) total ozone and not extreme days (NEOs – white area). GTO is an abbreviation for Grouped Total ozone Observations. Colored circles show global modes distinguished by the following indices: ENSO (light blue), NAO (orange), volcanic eruptions of Gunung Agung, El Chichón and Mt. Pinatubo (purple), polar vortex ozone loss (grey) and increasing/decreasing long-term development of ODS (black). Grey vertical bar denotes to a large fraction of missing data in 1929–1932. For references of atmospheric indices and other atmospheric quantities used see Part 2.

Table 2. Ozone mini-holes (threshold of 225 DU applied – following Brönnimann and Hood, 2003) observed over Arosa, Switzerland, in 1927–2008. Events marked with (*) indicate mini-holes that are found also following the stronger limiting definition of Bjokov and Balis (2001) (threshold of 220 DU).

year	month	day	total ozone [DU]	
1942	10	3	224	
1978	12	9	224	
1988	2	1	224	
1990	2	8	221	
1992	11	11	222	
1992	11	23	217	*
1992	12	2	215	*
1994	11	14	220	
1994	11	15	212	*
1996	11	7	221	
1998	1	1	220	
1999	1	6	223	
1999	12	1	210	*
1999	12	2	211	*
2000	11	29	198	*

4 Conclusions

Our analysis shows that a Gaussian distribution is inadequate to describe extremes in the daily mean values of the Arosa total ozone time series, because of a high frequency of low and high values, i.e. heavy tails. In this study we therefore apply tools from extreme value theory for the first time in the field of stratospheric ozone research. The use of the Generalized Pareto Distribution (GPD) is justified by arguments on asymptotic behaviour of data series (Pickands, 1975) for modelling observations above (below) a sufficiently high (low) threshold. Several tools for selecting appropriate thresholds have been applied on the Arosa ozone

time series (see Sects. 2.4 and 3.2). The comparison with the frequency distribution of the measured total ozone shows that these tools are appropriate to identify extreme values in the Arosa ozone data set. In order to model the seasonal variation a moving threshold (on a daily basis) was used. We show that the difference in mean values for the periods 1927–1959 and 1960–2008 is larger than the difference between the threshold values for ELOs and EHOs (see Sect. 3.2).

This paper also compares the results of extreme value theory with earlier publications dealing with ozone mini-holes. We show that the GPD-model fits the tail of the data and is suitable to estimate the probability of an ELO to describe mini-holes as used in previous papers.

The analysis of the observed frequency of extreme events shows a strong increase in the number of days with extreme low total ozone and a strong decrease (both about 5%) in days with extreme high total ozone during the last decades. Overall the application of extreme value theory provides new insights on the time series properties. As shown in Sect. 3.5, “fingerprints” of major atmospheric processes (e.g. ENSO, NAO, volcanic eruptions, and contributions of polar vortex ozone loss) can be clearly identified in the frequency distribution of ELOs and EHOs. In an accompanying paper (Part 2) we discuss the results of this type of analysis in terms of atmospheric processes as well as the influence of extremes on mean values, trends and long-term changes.

Supplementary material related to this article is available online at:

<http://www.atmos-chem-phys.net/10/10021/2010/acp-10-10021-2010-supplement.pdf>.

Acknowledgements. This work was performed within the project EXTREMES: “Spatial extremes and environmental sustainability: Statistical methods and applications in geophysics and the environment”, funded by the Competence Centre for the Environment and Sustainability (CCES) within the ETH-domain in Switzerland. The authors thank Herbert Schill for valuable discussion on the Arosa

ozone data. The authors are grateful to Robert Lund and one anonymous referee for useful comments during the discussion phase in ACPD. Further the authors are grateful to another anonymous referee for comments during the revision phase of the article.

The Software R (R Development Core Team, <http://R-project.org>) was used for the statistical analysis.

Edited by: M. Dameris

References

- Anton, M., Serrano, A., Cancillo, M. L., and Garcia, J. A.: Total ozone and solar erythemal irradiance in southwestern Spain: Day-to-day variability and extreme episodes, *Geophys. Res. Lett.*, 35, L20804, [10.1029/2008gl035290](https://doi.org/10.1029/2008gl035290), 2008.
- Appenzeller, C., Weiss, A. K., and Staehelin, J.: North Atlantic oscillation modulates total ozone winter trends, *Geophys. Res. Lett.*, 27, 1131–1134, 2000.
- Bojkov, R. D. and Balis, D. S.: Characteristics of episodes with extremely low ozone values in the northern middle latitudes 1957–2000, *Ann. Geophys.*, 19, 797–807, [doi:10.5194/angeo-19-797-2001](https://doi.org/10.5194/angeo-19-797-2001), 2001.
- Brönnimann, S. and Hood, L. L.: Frequency of low-ozone events over northwestern Europe in 1952–1963 and 1990–2000, *Geophys. Res. Lett.*, 30, ASC8-1-5, [doi:10.1029/2003gl018431](https://doi.org/10.1029/2003gl018431), 2003.
- Brönnimann, S., Luterbacher, J., Staehelin, J., Svendby, T. M., Hansen, G., and Svenoe, T.: Extreme climate of the global troposphere and stratosphere in 1940–42 related to El Niño, *Nature*, 431, 971–974, [doi:10.1038/nature02982](https://doi.org/10.1038/nature02982), 2004.
- Cleveland, R. B., Cleveland, W. S., McRae J. A., and Terpenning, I.: STL: A Seasonal-Trend Decomposition Procedure Based on Loess, *J. Official Stat.*, 6, 3–73, 1990.
- Cleveland, R. B., Cleveland, W. S., McRae J. A., and Terpenning I.: STL: A Seasonal-Trend Decomposition Procedure Based on Loess, *J. Official Stat.*, 6(1), 3–73, 1990.
- Cleveland, W. S.: Robust locally weighted regression and smoothing scatterplots, *J. Am. Stat. Assoc.*, 74, 829–836, 1979.
- Cleveland, W. S. and Devlin, S. J.: Locally weighted regression – an approach to regression analysis by local fitting, *J. Am. Stat. Assoc.*, 83, 596–610, 1988.
- Coles, S.: An introduction to statistical modeling of extreme values Springer series in Statistics, Springer, London, 2001.
- Coles, S. and Pericchi, L.: Anticipating catastrophes through extreme value modelling, *Journal of the Royal Statistical Society Series C-Applied Statistics*, 52, 405–416, 2003.
- Cooley, D., Nychka, D., and Naveau, P.: Bayesian spatial modeling of extreme precipitation return levels, *Journal of the American Statistical Association*, 102, 824–840, [doi:10.1198/016214506000000780](https://doi.org/10.1198/016214506000000780), 2007.
- Davison, A. C. and Smith, R. L.: Models for exceedances over high thresholds (with Discussion), *Journal of the Royal Statistical Society Series B*, 52, 393–442, 1990.
- Fawcett, L.: Statistical methodology for the estimation of environmental extremes, PhD-Thesis, University of Newcastle, Newcastle upon Tyne, 2005.
- Fawcett, L. and Walshaw, D.: A hierarchical model for extreme wind speeds, *Appl. Stat.*, 55, 631–646, 2006.
- Iwao, K. and Hirooka, T.: Dynamical quantifications of ozone mini-hole formation in both hemispheres, *J. Geophys. Res.-Atmos.*, 111, D02104, [doi:10.1029/2005jd006333](https://doi.org/10.1029/2005jd006333), 2006.
- James, P. M.: A climatology of ozone mini-holes over the northern hemisphere, *Int. J. Climatol.*, 18, 1287–1303, 1998.
- Koch, G., Wernli, H., Schwierz, C., Staehelin, J., and Peter, T.: A composite study on the structure and formation of ozone mini-holes and minihighs over central Europe, *Geophys. Res. Lett.*, 32, L12810, [doi:10.1029/2004gl022062](https://doi.org/10.1029/2004gl022062), 2005.
- Krzyżscin, J. W.: Long-term changes in ozone mini-hole event frequency over the Northern Hemisphere derived from ground-based measurements, *Int. J. Climatol.*, 22, 1425–1439, [doi:10.1002/joc.812](https://doi.org/10.1002/joc.812), 2002.
- Lund, R., Hurd, H., Bloomfield, P., and Smith, R.: Climatological time series with periodic correlation, *J. Climate*, 8, 2787–2809, 1995.
- Pickands, J.: Statistical-inference using extreme order statistics, *Ann. Stat.*, 3, 119–131, 1975.
- Ribatet, M.: POT: Modelling Peaks Over a Threshold, *R News*, 7, 3, 2007.
- Ribatet, M., Ouarda, T., Sauquet, E., and Gresillon, J. M.: Modeling all exceedances above a threshold using an extremal dependence structure: Inferences on several flood characteristics, *Water Resour. Res.*, 45, W03407, [doi:10.1029/2007wr006322](https://doi.org/10.1029/2007wr006322), 2009.
- Rieder, H. E., Staehelin, J., Maeder, J. A., Peter, T., Ribatet, M., Davison, A. C., Stübi, R., Weihs, P., and Holawe, F.: Extreme events in total ozone over Arosa – Part 2: Fingerprints of atmospheric dynamics and chemistry and effects on mean values and long-term changes, *Atmos. Chem. Phys.*, 10, 10033–10045, [doi:10.5194/acp-10-10033-2010](https://doi.org/10.5194/acp-10-10033-2010), 2010.
- Royston, P.: An extension of Shapiro and Wilk’s W test for normality to large samples, *Appl. Stat.*, 31, 115–124, 1982a.
- Royston, P.: Algorithm AS 181: The W test for Normality, *Appl. Stat.*, 31, 176–180, 1982b.
- Sassi, F., Kinnison, D., Boville, B. A., Garcia, R. R., and Roble, R.: Effect of El Niño-Southern Oscillation on the dynamical, thermal, and chemical structure of the middle atmosphere, *J. Geophys. Res.-Atmos.*, 109, D17108, [doi:10.1029/2003jd004434](https://doi.org/10.1029/2003jd004434), 2004.
- Sato, M., Hansen, J. E., McCormick, M. P., and Pollack, J. B.: Stratospheric aerosol optical depths, 1850–1990, *J. Geophys. Res.-Atmos.*, 98, 22987–22994, 1993.
- Staehelin, J.: Total ozone series of Arosa (Switzerland). Homogenization and data comparison, *J. Geophys. Res.*, 103, 5827–5841, 1998.
- Staehelin, J., Renaud, A., Bader, J., McPeters, R., Viatte, P., Hoegger, B., Bugnion, V., Giroud, M., and Schill, H.: Total ozone series at Arosa (Switzerland): Homogenization and data comparison, *J. Geophys. Res.-Atmos.*, 103, 5827–5841, 1998.
- Vogler, C., Brönnimann, S., Staehelin, J., and Griffin, R. E. M.: Dobson total ozone series of Oxford: Reevaluation and applications, *J. Geophys. Res.-Atmos.*, 112, D20116, [doi:10.1029/2007jd008894](https://doi.org/10.1029/2007jd008894), 2007.
- Wohltmann, I., Lehmann, R., Rex, M., Brunner, D., and Mader, J. A.: A process-oriented regression model for column ozone, *J. Geophys. Res.-Atmos.*, 112, D12304, [doi:10.1029/2006jd007573](https://doi.org/10.1029/2006jd007573), 2007.