DOI: 10.31401/WS.2025.proc.28

Basic Statistical Properties of the 10-years Ozone Daily Time Series from Stara Zagora

Werner R.[1], Petkov B.[2], Guineva V.[1], Bojilova R.[3], Raykova L.[1], Valev D.[1], Kirillov A.S.[4]

[1] Space Research and Technology Institute (SRTI) – Bulgarian Academy of Sciences, Stara Zagora Department, Bulgaria,

rolwer52@yahoo.co.uk

[2] Institute of Polar Sciences, National Research Council, Bologna, Italy

[3] National Institute of Geophysics, Geodesy and Geography (NIGGG) - Bulgarian Academy of Sciences, Sofia, Bulgaria

[4] Polar Geophysical Institute (PGI), Apatity, Russia

Abstract

In order to examine some important statistical parameters, the ozone series were detrended and seasonally adjusted. Then we determined the central moments as well as the skewness and kurtosis of the series. The distributions were analysed using histograms and Q-Q plots for different seasons. We divided the series into two groups with high and low ozone variations. The group with high variations includes the months from December to March (DJFM), while the group with lower variations includes the months from May to October (MJJAASO) with standard deviations of 36 DU and 19 DU, respectively. The reason for the high ozone variations is the frequent exchange of ozone-rich polar air masses with pure ozone air masses from the mid-latitudes and the tropics. The Q-Q plots show that the distributions for the two groups are close to normal distribution. On short time scales of a few days, the ozone residuals can be described by an AR (1) model. Long-term persistence was investigated using detrended fluctuation analysis (DFA). On scales ranging from a few days to a crossover of about 20 days, a Hurst coefficient of 0.93 was found, which characterises the series on these scales by strong persistence, far from chaotic behaviour but close to 1/f noise. From the crossover onwards, a Hurst coefficient of around 0.7 characterises the long-range correlation. The crossover achieved in the DFA of the seasonally adjusted and detrended total column ozone (TCO) indicates a change in the dominant processes on different time scales.

Keywords: statistic distribution, short-range correlation, Detrended Fluctuation Analysis

1. Introduction

Since the discovery of the Antarctic ozone hole, the mechanisms that cause ozone depletion have been intensively studied and the number of related publications is enormous. Here we will only refer to the reviews by Solomon [1999], Grewe [2007] and Chipperfield [2017]. Over the Antarctic, the lowest average ozone levels are reached in October. The October values decreased from the level 360 DU before 1980 by 120 DU in 1998-2000. The entry into force of the Montreal Protocol banning ozone depletion substances led to a decrease in the effective stratospheric chlorine loading (EESC) since the end of the 1990s. Since then, Antarctic ozone has not decreased any further. The same applies to ozone above the Arctic, although the decrease in ozone there is not as strong as in the Antarctic and the observed ozone depletion in March is only about 54 DU during 1998-2000 [Andersen and Knudsen, 2002]. At the Northern mid-latitudes (between latitudes of 30°N and 60°N) for the interval from 1979 up to the middle of 1990 the ozone decrease was about 5.5% [Harris et al., 2008]. That means the ozone loss was about 25 DU related to an averaged mean March total column ozone before 1979 of about 460 DU. Weber et al. [2022], using a multiple regression model based on different ozone annual satellite data sets, has found a significant downward trend of ozone depletion (from 1979 and

1995) for the latitude band 35°N – 60°N. But the observed trend of 0.5% per decade from 1995 up to 2020 is at the significance limit of 2σ. Fountoulakis et al. [2016] based on measurements with a Brewer spectralphotometer has been reported a TCO trend of 0.8%/decade for the time interval of 1994-2014 at Thessaloniki, a location near to Stara Zagora at a distance of only about 300 km in south-west direction. Werner et al. [2023a] did not find a significant trend in the ozone series from 1981 to 2022 based on the Sofia overpass data TM3DAM-MSR2. The increase in ozone following the observed trend reversal between 1995 and 1998 I s very slow and depends on latitude. The significance of the trends found is influenced by the correlation properties of the residuals of the ozone time series models. AR(1) models are generally used. However, the residuals also exhibit long-rang correlations, whereby the decline in the correlation function with increasing time delay can be described by a power law. Probability distributions and correlation properties are therefore among the fundamental characteristics of time series. They are essential, for example, when performing hypothesis tests. Significance tests of parameters in linear regressions, including trend parameters, are based on the assumption of a normal distribution of the residuals. The comparison of variances using the F-test also assumes a normal distribution. Skewness and kurtosis are used to assess whether the distribution of daily ozone values can be described by a normal distribution.

2. Data

This work is based on ground-based measurements of ultraviolet radiation intensity using a GUV 2511 global ultraviolet radiometer developed by Biospherical Instruments Inc. (BSI). Light, both direct sunlight and light scattered from the sky, falls on a Teflon entrance window and enters a hollow cylinder. This cavity contains seven photodiodes with filters in the UV range at wavelengths of 305, 313, 320, 340, 380, and 390 nm, each with an FWHM of approximately 10 nm, and a channel for detecting photosynthetically active radiation (PAR) in the visible range from 400 nm to 700 nm. The measurement process is controlled by a controller. The data files for each day are stored using a computer. Our device is stabilized at 50 °C. Further details about the device can be found in the brochure [BSI, 2008]. Biospheric Instruments subjects all GUV devices to radiation calibration in its optical laboratories.

To determine the total ozone in the atmosphere (TCO), we used the ratio of UV radiation at 340 nm, which is not affected by ozone, to radiation at 313 nm, which belongs to the absorption bands of ozone. In addition, the optical depth was determined using the ratio of the actual measured radiation at 380 nm and the radiation on a corresponding clear day. Look-up tables (LUT) (Stamnes et al., 1991) were created using the Madronich radiation transfer model for UV, known as the TUV model, which models the irradiance ratios and optical depths as a function of the zenith angle for different TCOs. The details of the algorithm, including the preprocessing procedures for TCO determination based on GUV UV irradiance measurements, are described in the publications [Werner et al., 2017 and 2018].

The stability of the results obtained with the GUV instrument and the developed algorithm was investigated by comparing them with the ozone values of the OMI instrument on the Aura satellite. It was found that, firstly, the absolute deviation of the successive four-year intervals 2015 –2018 and 2019–2022 using the OMI-Aura TCO as a reference increases slightly by 1.8 DU, and secondly, that the parameters of the linear regressions of both 4-year time series are not statistically different and do not differ from the pooled series. It was concluded that the temporal stability is acceptable [Werner et al., 2023b].

For the period from (February) 2015 to (March) 2025, we observed 3666 daily TCO values, which corresponds to 87% of the possible number. We only have no TCO data for about 13% of the days, partly for technical reasons and partly because of insufficient daylight due to heavy cloud cover. We used the ozone data from the OMI Aura satellite to complete our data series to about 97 %.

3. Ozone time series

The sequence of daily, monthly, seasonal, or annual ozone values for a location or average values for specific regions represent time series. Their components basically consist of a trend, a seasonal component, and a residual component. The trend T and the seasonal component S are often combined as a dynamically determined component. Time series models can be described by regression equations with time-dependent (possibly delayed) factors (predictors) x_t and delayed variables $y_{t-\tau}$ that describe their own history [Maddala, 1992]. For the trend, the factors X_t represent time itself, quadratic or higher exponents of time. By introducing dummy variables that describe a ramp function, a piecewise regression can be performed to detect structural breaks. Seasonality can be represented by sinusoidal factors. The statistical model for

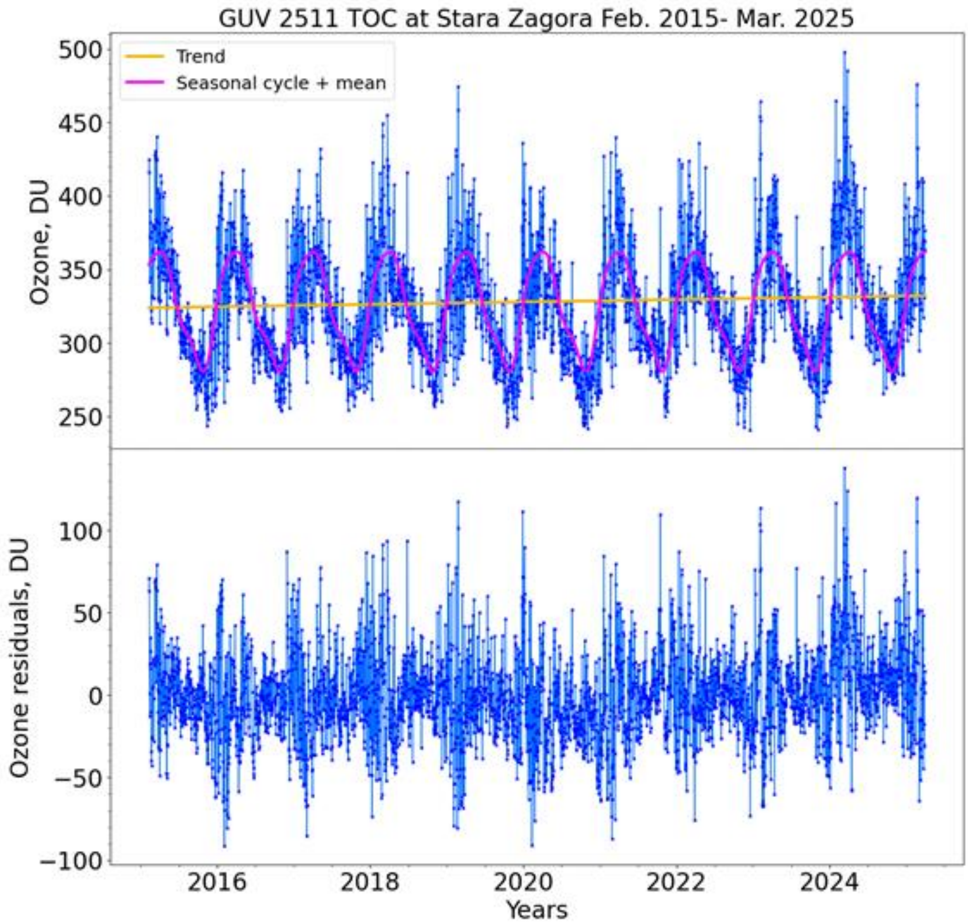

Figure 1. *The original TCO series (blue line and dots for measured values) in Stara Zagora for the period from February 2015 to March 2025, together with the annual seasonal cycle (magenta line) and the trend together with the mean value-(orange line), are shown in the upper field. The lower field shows the residuals. (The trend shown here was determined in a first step, and the seasonal cycle was calculated in a second step.)*

the time interval studied here (after the turning point, where a change in trend was observed) can therefore be represented as follows:

$$T_t = \mu + \alpha t + \sum_{i=1}^{k} \left[\beta_{2i-1} \sin\left(\frac{2\pi i t}{T_{per}}\right) + \beta_{2i} \cos\left(\frac{2\pi i t}{T_{per}}\right) \right] + \sum_{j=1}^{m} \gamma_j \ X_j + \varepsilon_t, \qquad (1)$$

where Xj are predictors as indices for solar activity and QBO and ε represent the residuals. However, the effects of QBO and solar flux on ozone in the mid-latitudes are very small (of the order of a few DU). For this reason, we limit our analysis of the ozone series to the seasonal component and the trend [see also Reinsel et al., 2002], whereby the mean value can be attributed to the trend component. We use three pairs of sinusoids for the seasonal component. The residuals ε_t must be stationary and normally distributed in order to perform tests, e.g., to perform the 2-sigma significance test of the regression coefficients in equation (1). Figure 1 shows the original TCO-Stara-Zagora series together with the seasonal and trend components in the upper panel and the residuals ε_t in the lower panel. The residual series shows clear non-stationarity. During the winter season, variability is much higher than during the rest of the year. This is caused by the frequent change of polar air masses with higher ozone content and air masses of tropical origin with lower TCO. In an earlier study, we showed through piecewise regression that the standard deviation in the winter months is about twice as high as in the other seasons [Werner, 2021]. The transition from the period with higher standard deviation to the period with lower standard deviation is related to the collapse of the vortex due to a sequence of minor and/or major sudden stratospheric warmings.

4. Standard deviation, skewness and kurtosis of the TCO residuals

The standard deviation, skewness, and kurtosis are non-Gaussian statistics that can be used to characterize time series. They are defined based on central moments of second, third, and fourth order. The skewness of single-peaked (unimodal) probability distributions describes the deviation from symmetry, and the kurtosis describes the sharpness of the peak.

To estimate the values of the standard deviation *std*, skewness *sk*, and kurtosis *k*, as well as their standard errors (*SE*), we use the traditional equations. Alternatively, the traditional statistics for skewness and kurtosis are multiplied by a function of the number of observations [Wright and Herrington, 2011]. These functions converge to 1 with $n \to \infty$. (In our work, n is of the order of 1000). The traditional equations are as follows [ibid.]:

=
1

$$sk = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^3}}, \qquad SE(sk)=\sqrt{\frac{6(n-2)}{(n+1)(n+3)}}, \qquad for \;\; n \gg 1: \qquad SE(sk) = \sqrt{\frac{6}{n}}, \qquad (2)$$

1

i=1nxi−x2, Confidence interval of SE (std): ixi−x2χn;1−α/22;ixi−x2χn;α/22 ,

$$k = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4}{\sqrt{\left(\frac{1}{n}\sum_{i=1}(x_i-\bar{x})^2\right)^4}}, \qquad SE(k) = \sqrt{\frac{24\,n\,(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}}, \qquad for \;\; n \gg 1: \qquad SE(k) = \sqrt{\frac{24}{n}},$$

=
1

where $\bar{x}$ is the mean of the observed values x_i. Because the values of the χ_n^2 distribution for the confidence levels 1-α/2 and α/2 are different, the confidences interval for the standard error of the standard deviation is not symmetrical and its lower and upper limits are given. In the statistic literature, the abbreviation g_1 and g_2 are often used for *sk* and *k*. An approximation of the standard errors of *sk* and *k* is given for observation numbers that are much larger than one. It can be seen, that for $n \gg 1$, *SE*(*k*) is twice as large as *SE*(*sk*). Instead of *k*, the excess kurtosis $k_e = k - 3$ is often given. In the case of a normal distribution, *sk* and k_e are equal to zero, which means that the distribution is symmetrical and not peaked. For sk values between -0.5 and 0.5, a normal distribution is assumed; if k_e is between -1 and 1, the distribution is well normally distributed [Hatem et al., 2022].

Bashenov et al. [2019] analysed the TCO series from 1994 to 2017 for the Tomsk station (56.50°N, 84.97°E). The authors found minimal relative standard deviations in July and August. Standard deviation, skewness, and kurtosis were used to classify the long-term TCO series from Arosa (1927–2000, 46.8N, 9.7E), Lerwick (1957–1999, 60.1°N, 1.2°W), and Camborn (1957–1999, 50.2°N, 5.3°W) [Tuomi et al., 2001]. The standard deviations at all three stations show greater variability during the winter-spring season. In order to compare the seasonal character of *std*, *sk* and k_e, we calculated these measures together with their standard error *SE*. The results are summarised in Table 1. The standard deviation in winter is

Table 1: *The number of observations n, the standard deviation std, the skewness sk, the excess kurtosis* k_e *and their standard errors calculated for the winter DJF, the spring MAM, the summer JJA and fall SON seasons and the same for the winter/early spring, here called vortex season, DJFM and for late spring to the autumn season, here called non-vortex season, MJJASO for all observations with and without outliers.*

	DJF	MAM	JJA	SON	DJFM	MJJASO	MJJASO outl. del.
n	887	923	879	867	1216	1765	1758
std in DU	36.5	29.8	17.2	19.2	36.5	18.5	17.9
1σ-error in DU	-1.6/1.8	-1.3/1.4	-0.8/0.8	-0.9/0.9	-1.4/1.5	-0.6/0.6	-0.6/0.6
sk	0.409	0.506	0.602	0.718	0.425	0.488	0.185
1σ-error	0.082	0.080	0.082	0.083	0.070	0.058	0.058
k_e	-0,047	0.726	1.666	1.942	0.046	1.707	0.435
1σ-error	0.163	0.160	0.164	0.165	0.140	0.116	0.116

approximately twice as high as in summer and autumn. This is in good agreement with the results obtained from the investigation of ozone variance in the Sofia overpass data TM3DAM-MSR2 from 1981 to 2018 based on the iterative method of cumulative sums of squares. It was found that the structure of the TCO series has two breakpoints each year, one near the development of the vortex and one related to the vortex break, and it was concluded that the Sofia overpass ozone series is highly heteroscedastic [Werner et al., 2021]. As it can be seen from Table 1, $sk = 0.409$ and $k_e = -0.047$ for the winter season, so the distribution is almost normal. In the JJA and SON seasons, the kurtosis is significantly greater than zero, but the characteristics are very similar, and the assumption that they come from the same population is obvious.

We have divided the GUV-TCO series into two parts. One part in which the polar vortex (vortex season, DJFM) develops, and another part in which no vortex is observed (non-vortex season, MJJASO). The extended winter season DJFM has almost the same characteristics as the winter season DJF, and the characteristics of the non-vortex season, which combines the seasons JJA and SON, are comparable to those of the individual seasons. The high values of *sk* and k_e for the summer and autumn seasons as well as for the non-vortex season indicate a non-Gaussian distribution of the TCO residuals for these time intervals. The months of April and November are considered transition months from one season to the next one.

5. Histogramms and Q-Q plots

Numerous statistical tests have been developed to verify normal distribution. They include the Kolmogorov-Smirnov test, the Liliefors test, the Shapiro-Wilk test and the Anderson-Darling test. With large sample sizes, even a small deviation from normal distribution would yield significant results [Öztuna et a.l, 2006, Ghasemi and Zahediasl, 2012]. To clarify our conclusions about the distributions – non-Gaussian or Gaussian – we have created histograms and Q-Q plots for the ozone residuals distributions and all seasons examined here. In Q-Q plots, the empirical or observed sample quantiles are plotted above the theoretical quantiles.

In an ideal Gaussian distribution, the Q-Q plot is identical to the bisecting line. The theoretical probabilities were estimated here with $p = i/(n+1)$, where $i = 1 \ldots n$. The graphs of the resulting histograms and Q-Q plots are shown in Figure 2. The histogram for the winter season shows a slight tail on the right side, which corresponds to the positive skewness. The tails of the distributions for the other seasons are somewhat heavier, which is caused by outliers, as documented in the Q-Q plots for sample quantiles greater than about three. The middle parts of the Q-Q plots are very close to the bisecting line. As expected, the kurtosis changes significantly to smaller values after the outliers are removed. The characteristics s_k and k_e for the vortex season DJFM and for the no vortex season MJJASO, for which outliers have been removed, indicate that a Gaussian distribution can be considered as acceptable. This impression is confirmed by examining the corresponding Q-Q plots (see Figure 3). The advantage of using combined seasonal ozone values lies in the larger number of observations. Of course, outliers must be taken into account when analysing extreme ozone value distributions.

6. ACF – short-range correlation

Residual regre-ssion equations, such as [1], are often autocorrelated. This means that in the case of first-order autoregression, the residuals can be written as follows:

$$\varepsilon_t = \varphi_1 \varepsilon_{t-1} + u_t, \tag{3}$$

where u_t is normally distributed. Whether the residuals can be described by a first-order or higher-order autoregression model or by a more general autoregression model can be investigated using the autocorrelation and partial autocorrelation functions of the residuals. Our colleagues in the Geophysical department of the National Institute of Geophysics, Geodesy and Geography have used combined satellite data from 1998 to 2008 for the Sofia location to create the autocorrelation function of the ozone residuals and have determined an exponential decay with a time constant of approximately 2.5 days [Total_ozone]. Furthermore, the PCAF have only one coefficient r_{11}, which is equal to φ_1. Here we go one step further. In addition to the autocorrelation function, we analyse also the partial autocorrelation function (PCAF). In an ideal AR(1) process, the autocorrelation coefficients ρ_k to a lag k are related to the autoregression coefficient φ_1 with the equation $\rho_k = \varphi_1^k$ being valid. The estimated ACF for k=1 yields an autoregression coefficient of 0.61. The identity $\varphi_1^k = \exp(k * \ln \varphi_1)$ shows the exponential decay of the autocorrelation function for an AR(1) process. The estimated $\widehat{ACF}$ and $\widehat{PCF}$ of the ozone residuals are presented in Figure 4. The $\widehat{ACF}$ shows the exponential decay for the first four lags τ, which are determined

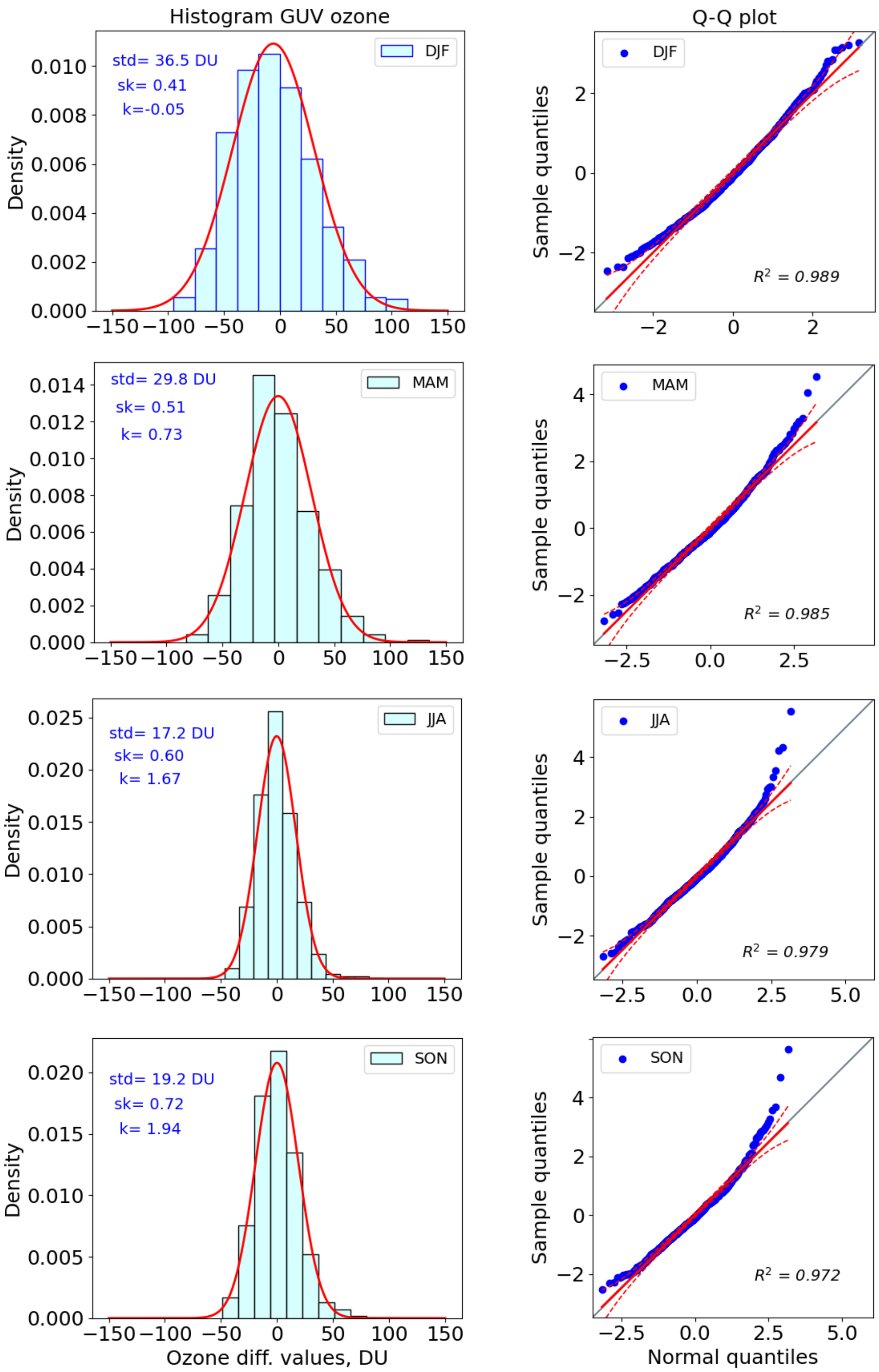

Figure 2. *Seasonal histograms of the ozone residuals together with their approximated Gausian distributions (red lines) are shown at left side and the corresponding Q-Q plots are presented at the right side of the panel. The red dashed lines are the upper and lower limits of the 0.95 confidence intervals. The graphics are displayed respectively for the DJF, MAM, JJA and SON from top to bottom.*

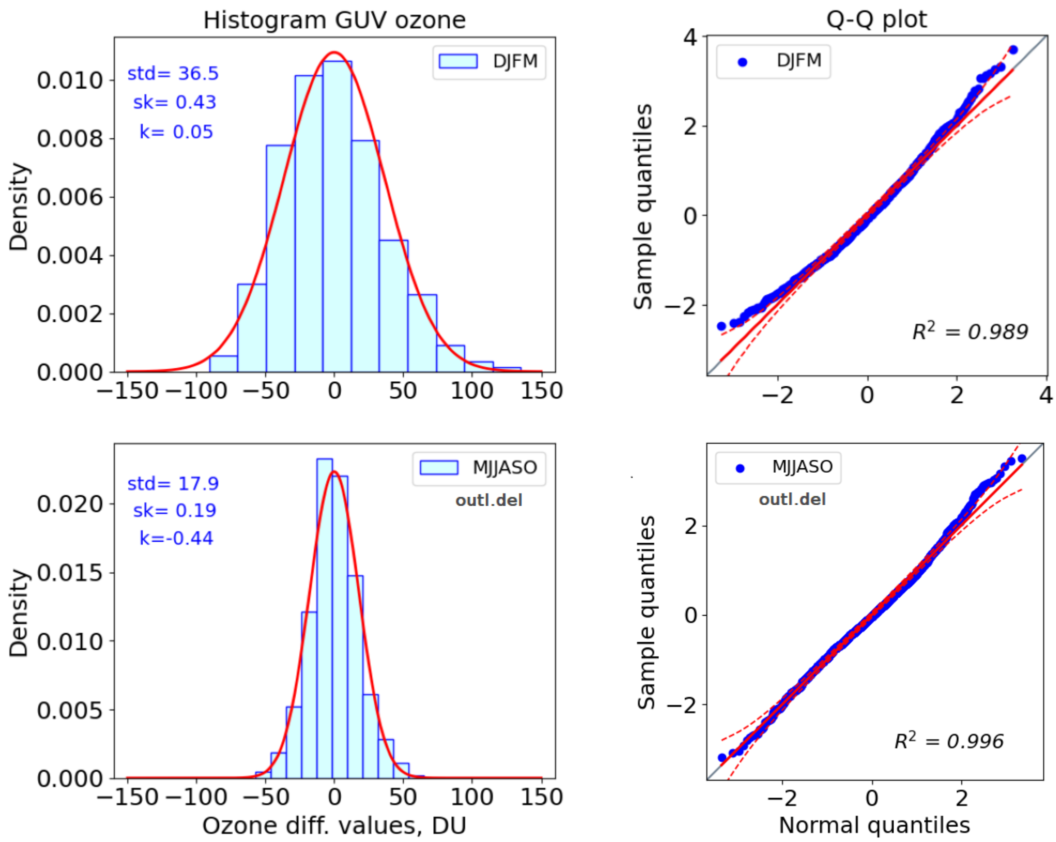

Figure 3. *The figure corresponds to Figure 2, but for the vortex season DJFM and for the non-vortex season MJJASO, where seven outliers were deleted for MJJASO (see Table 1).*

by a two-parameter approximation $\widehat{ACF} = c * exp\left(-\frac{\tau}{\tau_0}\right)$, drawn as red line in the inner figure at the left side of Figure 4. The time constant τ_0 was determined to be 2.70 days.

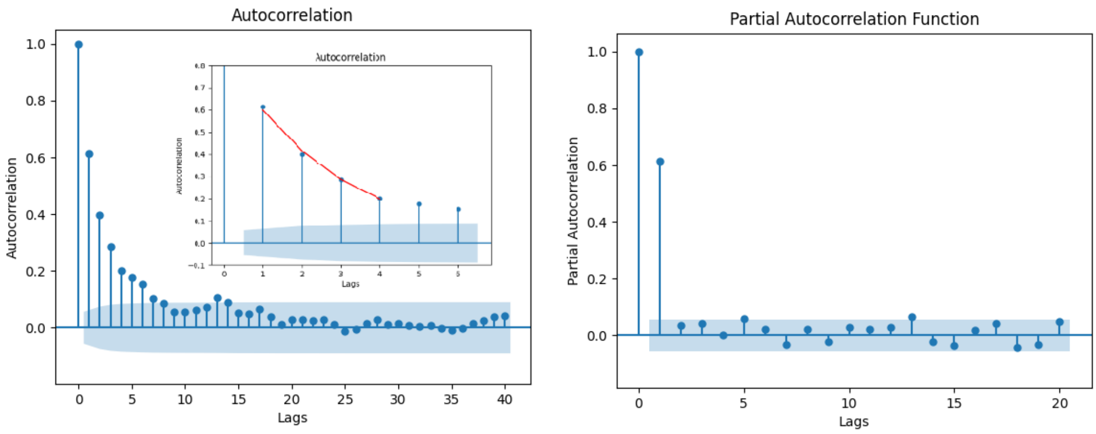

Figure 4. *The estimated autocorrelations r_k (left) and estimated partial autocorrelation function for the ozone residuals (right). The standard errors calculated by the Barlett equations, implemented in the Pythons ACF and PCF functions, for the corresponding lags are presented as shadowed regions.*

The estimated $\widehat{PCF}$ shows only one clear maximum at the lag 1, which characterises the order of the AR process. From this, we conclude that the ozone residual values can be describe by an AR(1) process. Ideally, an AR(1) process can be interpreted as a process with infinite memory. In practice, however, the ACF drops quickly and the memory is limited to only a few time steps. For the TCO series, this means that if the ozone has increased/decreased in the recent days, it can be assumed that it will continue to increase/decrease the next day. At greater lags, from four up to eight the autocorrelation coefficients decline more slowly and subsequently fall below the one-sigma error.

7. Detrended Fluctuation Analysis – Long persistence

The variation of correlation over time allows for interferences about the dynamics of complex physical processes. The direct calculation of autocorrelation for large lags is limited by the noise as demonstrated in the previous paragraph. A tool that is commonly used to investigate long persistence or long memory process is the Detrended Fluctuation Analysis (DFA) developed by Peng et al. [1994]. This method was further developed by Kantelhardt et al. [2002] into Multifractal Detrended Fluctuation Analysis (MF-DFA) to study multifractal processes. The results of the DFA enable the indirect determination of the decay of the autocorrelation function for large lags.

The DFA is determined in successive steps, starting with the calculation of the cumulative sum of the members of the time series, referred to as the profile or integrated anomaly. The series is then divided into N_s non-overlapping segment intervals of length s and the (local) linear trend is determined for each of these segments. (In the MF-DFA, polynomials of degree m are also permitted for this purpose). The squared values of the fluctuations in each segment are calculated by the sum of the mean square deviations of the detrended profile values.

The quadratic values of the fluctuations over all segments N_s are totalled and averaged. As not all values in the series are recorded during segmentation, the procedure is repeated from the end of the time series and the quadratic fluctuations are determined again. This results in a total of 2 N_s segments. The fluctuations as a function of the scales s result from their square root of the mean value of all 2 N_s fluctuations. If the data have a power law:

$$C(s) \sim s^{-\gamma} \text{ , } where \;\; 0 < \gamma < 1 \tag{4}$$

the fluctuation function also increases by a power law [Koscielny-Bunde et al., 1998] $F(s) \sim s^{\alpha}$. The exponent α is referred to as scaling exponent or correlation exponent and can be found as the slope of the F(s) in double logarithmic representation of F(s) over s. The relationship between α and γ is given by $\alpha = 1 - \frac{\gamma}{2}$. The correlation exponent is an estimate of the Hurst coefficient and determines the correlation properties of the series. For H = 0.5, the time series is not correlated and represents white noise. For scale intervals with H > 0.5, the series is broadly positive, the trends continue, and for H < 0.5 it is broadly negatively correlated, the trends reverse. The power spectrum function decays as: [see e.g. Varotsos, 2005]

$$S(f) \sim f^{-\beta} \text{ with } \beta = 2\alpha - 1. \tag{5}$$

[see e.g. Varotsos, 2005]. For α = 1 we obtain the power spectrum exponent β = 1 and *S(f)* ~ 1/f, corresponds to the so called *1/f* noise.

The mathematical equations of the DFA method are described in detail in Kantelhardt et al. [2002]. In recent decades, DFA and MF-DFA have been used to analyse time series in fields such as hydrology, biology, neuroscience and geophysics. Koscienlny-Bunde et al. [1998] analysed the long-term records of daily maximum temperatures (typically longer than 100 years) from 14 randomly selected surface stations around the world and determined a scaling exponent of 0.65.

Using DFA, Király found a latitudinal dependence of the scaling exponent for the daily mean temperature series [see citation in Kiss et al. 2007]. The exponents are of the order of 0.6 to 0.98, with the maximum near the equator and the minima in mid-latitudes in both hemispheres. Kiss et al. [2007] performed a DFA analysis for filtered and merged daily

Nimbus7 Earth Sample records (for each grid point between 65°S and 65°N latitudes in the Total Ozone Mapping Spectrometers (TOMS) and found a similar latitude dependence of the ozone scaling exponent as for the mean temperatures mentioned above with a maximum of 0.96 and a minimum of about 0.66. The authors conclude that the source of the long-range correlation lies in the atmospheric dynamics for both total ozone and temperature.

We applied DFA to analyse the long-range correlation of our total column ozone time series. As it is well known, non-stationarities such as cycles and trends can influence the resulting correlation.

The QBO effect on the TCO is important in the tropics and very small in mid-latitudes. Solar activity during the period of interest (2015-2024) is relatively low. At the beginning of 2015, the solar activity is in the decline phase after the maximum of 24th solar day. The F10.7 cm index averaged over 27 days decreases from about 131 solar flux units (sfu) to a minimum of 67 sfu after 2017 and starts to increase again after 2021, reaching its maximum of 252 sfu (OMNIWeb Results) in the second half of 2024.

The corresponding ozone response to solar activity is about 3.7 DU/100 sfu [Stolarski et al., 1991]. Reinsel et al. [2002] found a latitudinal dependence of the ozone response to solar variability of 6 to -2 DU/100 sfu. A response of 2 DU/100 sfu is given for the latitude band at 40°N. Calisesi and Matthes [2006] give an average value of 1-2 % total ozone change in the mid-latitudes from solar minimum to maximum. In the Ozone Assessment Report 2018, a response of -1 to 2 % of the total ozone anomaly is given for the latitude range 35° - 60°N. Knibbe et al. [2014] found no significant solar effect on the total ozone over the eastern Balkans with a confidence level of 99 %. Thus, we expect a mean ozone response of less than 3 DU over Stara Zagora for the period analysed here, and we do not include solar activity in our regression equation (1). This corresponds to the obtained results in Werner et al. [2023a], where a change from solar minimum to maximum of -2.5 DU to 5 DU for the time interval 1979 – 2022, which include the very strong solar cycles 21 and 22 was found. We only consider the linear trend (including the mean) and the seasonal cycle in the form of three sinusoids together in one equation. In this way, we obtain a linear trend of about 2.7 DU/decade. (We emphasis that this is not a long-term trend, as our ozone series is too short for such an analysis).

The DFA shows artificial deviations on small scales [Kantelhardt et al., 2001]. Therefore, we have to calculate the DFA for random series that have approximately the same length as our ozone series. DFA for such a random series is shown in Figure 5. As it was expected, the linear slope is 0.5. The shorter the scale length, i.e. shorter than 7-8 days, the more pronounced the deviation from the linear slope. Instead of correcting this effect with the correction factor function given in Kantelhardt et al., [2001], we do not interpret the DFA results for scales shorter than 7-8 days. The calculated DFA of our deseasonalised and detrended TCO is shown in Fig. 6. In our case, detrending has no influence on the resulting DFA, therefore no higher order DFA was calculated. The overall slope of the fluctuation function is 0.76 (see Figure 6). This is in good agreement with the investigated DFA of the daily deseasonalised ozone series from Arosa (1926-1970, 46.8°N, 9.7°E), from Camborn (1957-1999, 50.2°N, 5.3°W) and from Lerwick (1957-1999, 60.1°N, 1.2°W), where the Hurst coefficients obtained are 0.76, 0.76 and 0.77, respectively, with the rescaled range analysis [Tuomi et al., 2001]. For the entire Arosa series (1927-2000), H=0.80 was determined. The DFA of our TCO shows a crossover on a scale of about 20 days and a slope of 0.93 before and 0.67 after. Kiss et al. [2007] show the DFA of Arosa with better graphical resolution. We have derived a transition on a scale of about 30 days from this graph, with a slope of 0.90 before and 0.68 after it. This result agrees

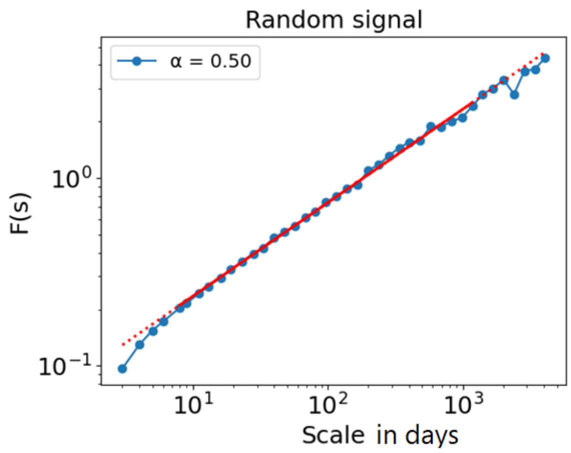

Figure 5. *DFA of a random signal. The slope of 0.5 is drawn by a continuous red line for the interval where a linear trend is observed. The dashed line extension is shown to outline the artificial deviations for small scales.*

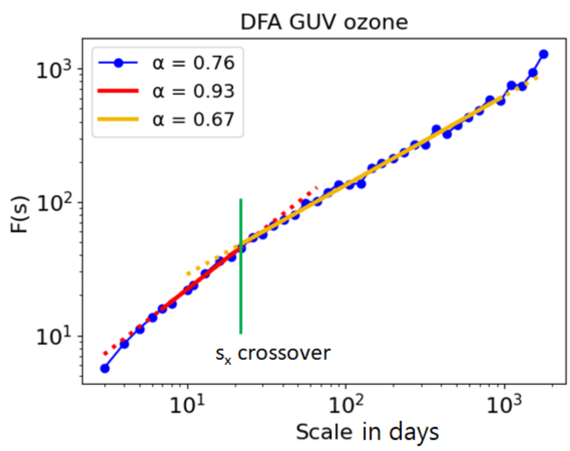

Figure 6. *DFA of the deseasonalized and detrended TCO at Stara Zagora. The location of the crossover is marked by a green vertical line. The slopes are drawn by continuous lines, which are extended to show their differences more clearly.*

very well with our findings. Kantelhardt et al. [2006] observed a transition on a scale of a few weeks in river discharge fluctuations and found that the time scale of the transition is similar to the period of planetary waves with typical periods of 2 to 30 days. Varotsos investigated the power-law correlations in the ozone column over Antarctica. He found an overlap in the DFA on a scale of about 10 days and concluded that this overlap illustrates the crucial role of planetary waves in the scaling characteristics of the spatio-temporal variability of the Antarctic ozone hole [Varotsos, 2005]. Kalamaras et al. [2017] found memory characteristics by H of 0.69, 0.67 and 0.69 for the daily mean, maximum and minimum temperature in Thessaloniki, respectively, and H=0.69 for the wind speed. It is well known that weather phenomena associated with tropospheric high or low pressure systems, which are reversible in the lower stratosphere, influence the TCO mainly through dynamical processes. Eicher et al. [2003] analyzed daily temperature records and determined a transition time of the order of 10 days, which is typical for a Großwetterlage (large-scale weather pattern) in the fluctuation function.

To test whether the non-stationarity caused by the different standard deviations during the vortex season (December–March) and non-vortex season (May–October) generated an artificial crossover at about 20 days, we calculated the DFA separately for the two seasons. The calculations were performed in two ways. Firstly, the DFA was determined individually for each segment and each year, and then the annual average was calculated. Secondly, the DFA was determined by concatenating (called "cut and stitching together", see Chen et al. [2002]) the seasonal parts, i.e. all Dec.-Mar. segments and all May-Oct. segments, correspondingly. The non-vortex season May-Oct. comprises 50% of a year and according to Cheng et al [2002] the scaling behaviour of the remaining part must be the same as for the full part. In fact, both series - the full series (Fig. 6) and of the stitching together series segments from May to October (Figure 7) - show a crossover s_x on approximately the same scale, and the slopes before and after the crossover are not noticeably different. For the series consistent of the sequences Dec. – Mar., stiching together, s_x is about 30 days and the correlation up to the crossover is also the same as for the full series. The long-range correlation after the crossover is somewhat lower. The comparison of the slopes after the crossover shows, that the

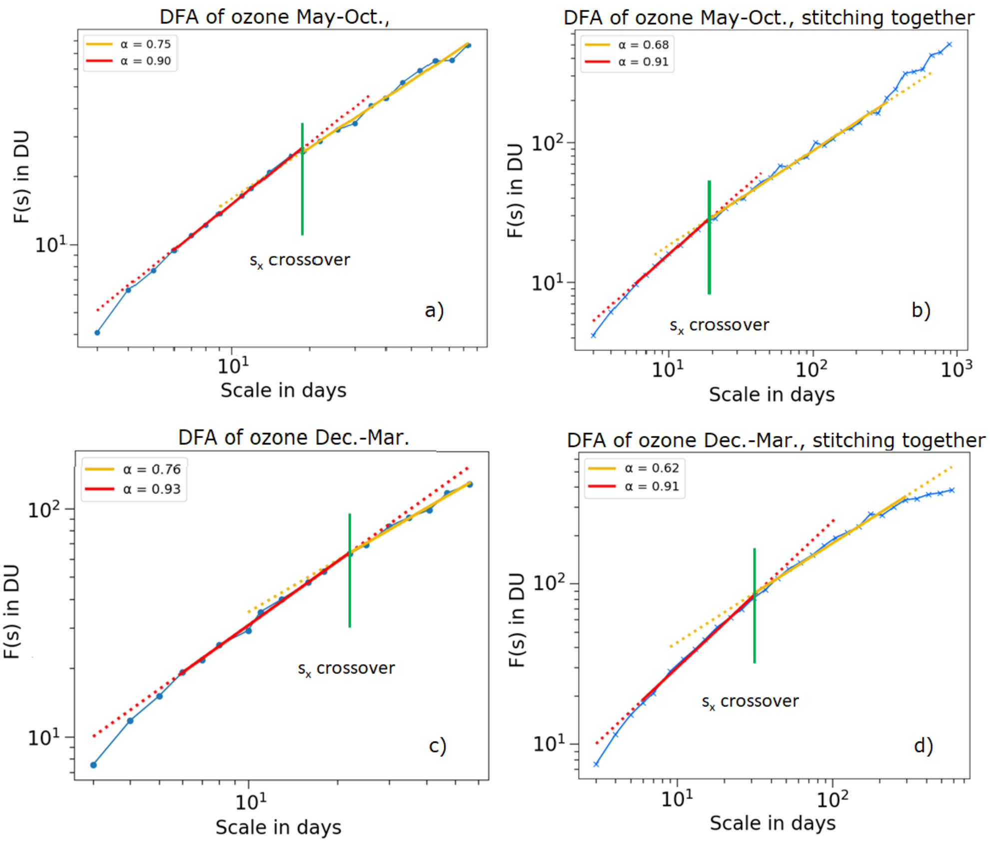

Figure 7. *Results of the DFA calculation. The left side shows the calculation of the DFA for each segment separately for each year with subsequent averaging (a and c) and the right side shows the DFA obtained by „cut and stitching together"(b and d). The upper panels show the DFA for the segments May-Oct. and the lower panels the DFA for the segments Dec.-March.*

slopes for the series from December to March and for the series from May to October, calculated by annual averaging, are significantly greater than the slope for the complete series. The reason for this may lie in the shortness of the remaining interval for determining the long-range correlation. To summarise, it can be said that the overlap occurred independently of the method used to calculate the DFA, so we conclude that the crossover is not an artificial effect.

The entire non-stationary TCO series consists of segments with two different standard deviations. These are the segments that cover the months from December to March and from May to October. The series part formed by stitching together the segments from December to March and the one formed by stitching together segments from March to October are stationary in themselves. The scaling behaviour of the non-vortex series of the segments from March to October (Figure 7) does not differ from that of the non-stationary TCO series (Figure 6). This corresponds very well to the results of Chen's simulations for non-stationary correlated signals ($\alpha > 0.5$) with segments characterised by two different standard deviations [Cheng et al., 2002].

8. Summary

The values for skewness and kurtosis as well as the Q-Q diagram confirm the normality of the ozone residuals for the winter season DJF. The characteristics standard deviation, skewness and excess of kurtosis of the summer season JJA and the autumn season SON are similar, and it is reasonable to assume that the ozone residuals originate from the same statistical sample.

The entire deseasonalised and detrended TCO series is non-stationary (heteroscedastic). We compiled samples from only two subsamples of the residuals. Namely, one for the season

December to March, when ozone in Stara Zagora is influenced by the alternation of polar air masses with tropical air, and the other for May to October. The absolute values of skewness and excess of kurtosis of the combined samples are less than 0.5 and their distributions can be considered as normal.

The DFA of the annual GUV ozone residuals shows that the series have a long time memory. A crossover was observed at scale s_x of about 20 days. Before s_x a Hurst coefficient of about 0.93 and after s_x a Hurst coefficient of about 0.68 was observed. The crossover is caused by various atmospheric processes. The fluctuations before the crossover are probably related to planetary waves and weather phenomena, while after the crossover the source of the long-range correlation is related to large-scale atmospheric dynamical processes.

Reference

Andersen, S.B., Knudsen, B.M. (2002). The influence of vortex ozone depletion on Arctic ozone trends, *Geophysical Research Letters*, Vol. 29, No. 21, pp. 9-1–9-4, DOI: 10.1029/2001GL014595.

Bazhenov, O.E., Elnikov, A.V., Sysoev, S.M. (2019). Total ozone content over Tomsk in 1994–2017: Results of statistical analysis, *Atmospheric and Oceanic Optics*, Vol. 32, No. 6, pp. 680–685, DOI: 10.1134/S1024856019060034.

Biospherical Instruments Inc. (2008). *Model GUV-2511 — Ground-Based UV / PAR Radiometer* (brochure).

Calisesi, Y., Matthes, K. (2006). The middle atmospheric ozone response to the 11-year solar cycle, *Space Science Reviews*, Vol. 125, pp. 273–286, DOI: 10.1007/s11214-006-9063-4.

Chen, Z., Ivanov, P.Ch., Hu, K., Stanley, H.E. (2002). Effect of nonstationarities on detrended fluctuation analysis, *Physical Review E*, Vol. 65, 041107, DOI: 10.1103/PhysRevE.65.041107.

Chipperfield, M.P. (2015). Global atmosphere – the Antarctic ozone hole, in: Harrison, R.M., Hester, R.E. (eds.) *Still Only One Earth: Progress in the 40 Years Since the First UN Conference on the Environment*, pp. 1–33, DOI: 10.1039/9781782622178-00001.

Eichner, J.F., Koscielny-Bunde, E., Bunde, A., Havlin, S., Schellnhuber, H.-J. (2003). Power-law persistence and trends in the atmosphere: A detailed study of long temperature records, *Physical Review E*, Vol. 68, 046133, DOI: 10.1103/PhysRevE.68.046133.

Fountoulakis, A., Bais, A., Fragkos, K., Meleti, C., Tourpali, K., Zempila, M.M. (2016). Short- and long-term variability of spectral solar UV irradiance at Thessaloniki, Greece: Effects of changes in aerosols, total ozone and clouds, *Atmospheric Chemistry and Physics*, Vol. 16, pp. 2493–2505, DOI: 10.5194/acp-16-2493-2016.

Ghasemi, A., Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians, *International Journal of Endocrinology and Metabolism*, Vol. 10, No. 2, pp. 486–489, DOI: 10.5812/ijem.350.

Grewe, V. (2007). Impact of climate variability on tropospheric ozone, *Science of the Total Environment*, Vol. 374, pp. 167–181, DOI: 10.1016/j.scitotenv.2007.01.032.

Harris, N.R.P., Kyrö, E., Staehelin, J., Brunner, D., Andersen, S.-B., Godin-Beekmann, S., Dhomse, S., Hadjinicolaou, P., Hansen, G., Isaksen, I., Jrrar, A., Karpetchko, A., Kivi, R., Knudsen, B., Križan, P., Lastovicka, J., Maeder, J., Orsolini, Y., Pyle, J.A., Rex, M., Vaníček, K., Weber, M., Wohltmann, I., Zanis, P., Zerefos, C. (2008). Ozone trends at northern mid- and high latitudes – a European perspective, *Annales Geophysicae*, Vol. 26, pp. 1207–1220, DOI: 10.5194/angeo-26-1207-2008.

Hatem, G., Zeidan, J., Goossens, M.M., Moreira, C. (2022). Normality testing methods and the importance of skewness and kurtosis in statistical analysis, *BAU Journal – Science and Technology*, Vol. 3, No. 2, Article 7, DOI: 10.54729/KTPE9512.

Kalamaras, N., Philippopoulos, K., Deligiorgi, D. (2017). Scaling properties of meteorological time series using detrended fluctuation analysis, in: Karacostas, T., Bais, A., Nastos, P. (eds.) *Perspectives on Atmospheric Sciences*, Springer Atmospheric Sciences, Springer, Cham, DOI: 10.1007/978-3-319-35095-0_78.

Kantelhardt, J.W., Koscielny-Bunde, E., Rego, H.H.A., Havlin, S., Bunde, A. (2001). Detecting long-range correlation with detrended fluctuation analysis, *Physica A: Statistical Mechanics and its Applications*, Vol. 295, pp. 441–454, DOI: 10.1016/S0378-4371(01)00144-3.

Kantelhardt, J.W., Zschiegner, S.A., Koscielny-Bunde, E., Havlin, S., Bunde, A., Stanley, H.E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series, *Physica A: Statistical Mechanics and its Applications*, Vol. 316, pp. 87–114, DOI: 10.1016/S0378-4371(02)01383-3.

Kantelhardt, J.W., Koscielny-Bunde, E., Rybski, D., Braun, P., Bunde, A., Havlin, S. (2006). Long-term persistence and multifractality of precipitation and river runoff records, *Journal of Geophysical Research: Atmospheres*, Vol. 111, D01106, DOI: 10.1029/2005JD005881.

Kiss, P., Müller, R., Jánosi, I.M. (2007). Long-range correlations of TOMS total ozone, *Nonlinear Processes in Geophysics*, Vol. 14, pp. 435–442, DOI: 10.5194/npg-14-435-2007.

Knibbe, J.S., van der A, R.J., de Laat, A.T.J. (2014). Spatial regression analysis on 32 years total column ozone data, *Atmospheric Chemistry and Physics Discussions*, Vol. 14, pp. 5323–5373, DOI: 10.5194/acpd-14-5323-2014.

Koscielny-Bunde, E., Bunde, A., Havlin, S., Roman, H.E., Goldreich, Y., Schellnhuber, H.-J. (1998). Indication of a universal persistence law governing atmospheric variability, *Physical Review Letters*, Vol. 81, No. 3, pp. 729–732, DOI: 10.1103/PhysRevLett.81.729.

Maddala, G.S. (1992). *Introduction to Econometrics* (2nd ed.). Macmillan Publishing.

Öztuna, D., Elhan, A.H., Tüccar, E. (2006). Investigation of four different normality tests in terms of type I error rate and power under different distributions, *Turkish Journal of Medical Sciences*, Vol. 36, No. 3, pp. 171–176.

Peng, C.K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E., Goldberger, A.L. (1994). Mosaic organization of DNA nucleotides, *Physical Review E*, Vol. 49, pp. 1685–1689, DOI: 10.1103/PhysRevE.49.1685.

Reinsel, G.C., Weatherhead, E., Tiao, G.C., Miller, A.J., Nagatani, R.M., Wuebbles, D.J., Flynn, L.E. (2002). On detection of turnaround and recovery in trend for ozone, *Journal of Geophysical Research: Atmospheres*, Vol. 107, DOI: 10.1029/2001JD000500.

Solomon, S. (1999). Stratospheric ozone depletion: A review of concepts and history, *Reviews of Geophysics*, Vol. 37, No. 3, pp. 275–316, DOI: 10.1029/1999RG900008.

Stamnes, K., Slusser, J., Bowen, M. (1991). Derivation of total ozone abundance and cloud effects from spectral irradiance measurements, *Applied Optics*, Vol. 30, No. 30, p. 4418, DOI: 10.1364/AO.30.004418.

Stolarski, R.S., Bloomfield, P., McPeters, R.D., Herman, J.R. (1991). Total ozone trends deduced from Nimbus-7 TOMS data, *Geophysical Research Letters*, Vol. 18, No. 6, pp. 1015–1018, DOI: 10.1029/91GL01302.

Tuomi, R., Syroka, J., Barnes, C., Lewis, P. (2001). Robust non-Gaussian and long-range correlation of total ozone, *Atmospheric Science Letters*, Vol. 2, Nos. 1–4, pp. 94–103, DOI: 10.1006/asle.2001.0042.

Varotsos, C. (2005). Power-law correlations in column ozone of Antarctica, *International Journal of Remote Sensing*, Vol. 26, No. 16, pp. 3333–3342, DOI: 10.1080/01431160500076111.

Weber, M., Arosio, C., Coldewey-Egbers, M., Fioletov, V.E., Frith, S.M., Wild, J.D., Tourpali, K., Burrows, J.P., Loyola, D. (2022). Global total ozone recovery trends attributed to ozone-depleting substance changes derived from five merged ozone datasets, *Atmospheric Chemistry and Physics*, Vol. 22, pp. 6843–6859, DOI: 10.5194/acp-22-6843-2022.

Werner, R., Guineva, V., Valev, D., Bojilova, R., Kirillov, A. (2023a). Long-time trend in total ozone column by MRS2 overpass data for Sofia for 1979–2022 based on ordinary and weighted linear regression, *15th Workshop “Solar Influences on the Magnetosphere, Ionosphere, and Atmosphere”* (PowerPoint presentation).

Werner, R., Petkov, B., Guineva, V., Kirillov, A., Atanassov, A., Bojilova, R., Valev, D., Raykova, L. (2023b). Verification of the long-term stability of ozone measurements with the GUV-2511 instrument in Stara Zagora, *Proceedings of the Nineteenth International Scientific Conference “SPACE, ECOLOGY, SAFETY”*, 24–26 October, Sofia, Bulgaria, pp. 58–62, p-ISSN 2603-3313.

Werner, R., Petkov, B., Valev, D., Guineva, V., Atanassov, A., Danov, D., Bojilova, R., Kirillov, A. (2021). Study of the Sofia total ozone content time series variance, available at: https://spaceclimate.bas.bg/SW/assets/pdf/Primorsko_Werner_1.pdf

Werner, R., Valev, D., Guineva, V., Kirillov, A., Petkov, B. (2018). Determination of the UV-index using measurements performed by the GUV-2511 instrument at Stara Zagora, *Proceedings of the Fourteenth International Scientific Conference “SPACE, ECOLOGY, SAFETY”*, 7–9 November, Sofia, Bulgaria, pp. 101–105, p-ISSN 2603-3313.

Wright, D.B., Herrington, J.A. (2011). Problematic standard errors and confidence intervals for skewness and kurtosis, *Behavior Research Methods*, Vol. 43, pp. 8–17, DOI: 10.3758/s13428-010-0044-x.