JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 115, D16306, doi:10.1029/2009JD013105, 2010

On the statistical modeling of persistence in total ozone anomalies

D. I. Vyushin,[1] T. G. Shepherd,[1] and V. E. Fioletov[2]

Received 29 August 2009; revised 13 April 2010; accepted 16 April 2010; published 21 August 2010.

[1] Geophysical time series sometimes exhibit serial correlations that are stronger than can be captured by the commonly used first-order autoregressive model. In this study we demonstrate that a power law statistical model serves as a useful upper bound for the persistence of total ozone anomalies on monthly to interannual timescales. Such a model is usually characterized by the Hurst exponent. We show that the estimation of the Hurst exponent in time series of total ozone is sensitive to various choices made in the statistical analysis, especially whether and how the deterministic (including periodic) signals are filtered from the time series, and the frequency range over which the estimation is made. In particular, care must be taken to ensure that the estimate of the Hurst exponent accurately represents the low-frequency limit of the spectrum, which is the part that is relevant to long-term correlations and the uncertainty of estimated trends. Otherwise, spurious results can be obtained. Based on this analysis, and using an updated equivalent effective stratospheric chlorine (EESC) function, we predict that an increase in total ozone attributable to EESC should be detectable at the 95% confidence level by 2015 at the latest in southern midlatitudes, and by 2020–2025 at the latest over 30°–45°N, with the time to detection increasing rapidly with latitude north of this range.

Citation: Vyushin, D. I., T. G. Shepherd, and V. E. Fioletov (2010), On the statistical modeling of persistence in total ozone anomalies, *J. Geophys. Res.*, *115*, D16306, doi:10.1029/2009JD013105.

1. Introduction

[2] When determining confidence intervals for long-term trends, it is necessary to take account of serial correlations of the residuals. It is convenient to study the serial correlations by analyzing the power spectrum of the residuals. It is also convenient to approximate this power spectrum with a parsimonious statistical model characterized by only a few parameters. Because the lowest-frequency part of the spectrum is always unobserved, it is both necessary and important to make an assumption about the spectral behavior in that range. Typically this assumption is implicitly made by the choice of model used to fit the spectrum. For the calculation of trends in total ozone, it is standard practice to model the residuals using a first-order autoregressive, or AR(1), statistical model [e.g., *Harris et al.*, 1998], which assumes that the serial correlations decay exponentially with time and equivalently that the power spectrum saturates to a constant in the low frequencies. Alternatively it can be assumed that the residuals exhibit what are called 'long-range correlations' (LRC) where the correlations decay only algebraically (i.e. more slowly) with time and the power spectrum increases by a power law in the low-frequency limit. In such cases, random trends over finite time intervals can be noticeably larger than predicted by the AR(1) model, which means that the confidence intervals on derived trends need to be widened, and the time required to detect a change in behavior (such as ozone recovery) lengthened. *Vyushin et al.* [2007, hereafter V07] analyzed total ozone time series over 1979–2005 and found that, while there was no evidence of LRC behavior in Southern Hemisphere (SH) middle and high subpolar latitudes, there was some evidence of such behavior in the Northern Hemisphere (NH). Taking account of LRC behavior widens the uncertainty of the magnitude of anthropogenic ozone depletion and lengthens the time required to detect ozone recovery by about a factor of 1.5 in the NH, and confirms the finding of *Weatherhead et al.* [2000] based on an AR(1) model that the high subpolar latitudes of the SH are the best place to detect recovery in total ozone.

[3] In this article we first assess the validity of the two above mentioned assumptions for the total ozone power spectrum using a novel approach based on the behavior of serial correlations under temporal aggregation. We show that the power law assumption can be used as an upper bound for total ozone persistence and therefore its consideration leads to an upper bound for trend uncertainty. The latter can be evaluated through an estimation of the Hurst exponent H, where the power spectrum $S(\lambda)$ is assumed to scale like $S(\lambda) \sim \lambda^{1-2H}$ as the frequency $\lambda \to 0$ [see, e.g., *Taqqu*, 2002]. However, as emphasized by various authors [V07; *Jánosi and Müller*, 2005; *Marković and Koch*, 2005; *Vyushin and Kushner*, 2009], the Hurst exponent estimation

[1]Department of Physics, University of Toronto, Toronto, Ontario, Canada.
[2]Environment Canada, Toronto, Ontario, Canada.

0148-0227/10/2009JD013105

needs to be handled with care as it can be affected by artifacts associated with the data processing. V07 and *Vyushin and Kushner* [2009] noted especially the possible sensitivity of the estimated H to the frequency bandwidth over which H is estimated, and to the presence of periodicities in the time series which can artificially increase H. *Kiss et al.* [2007, hereafter K07] analyzed long-range correlations in total ozone by means of detrended fluctuation analysis of the third order (DFA3 [*Kantelhardt et al.*, 2001]). The latitudinal distribution of the Hurst exponent estimate obtained in this manner has local maxima over high latitudes and the equator and local minima over midlatitudes. This distribution is quite different from the result found by V07. The difference was noted by K07, but not explained. Here we explain the difference in the two results, which has important implications for the detection of ozone recovery. Apart from clarifying the results for total ozone, our analysis also has implications for the estimation of the Hurst exponent in other geophysical time series, as it highlights the possible sensitivities of these estimates to choices made in the data analysis.

[4] We also update V07's analysis of the early stages of ozone recovery by considering a longer time series and an improved representation of EESC.

2. Data

[5] The merged satellite data set used here is prepared by NASA and combines version 8 of TOMS, SBUV, and OMI total ozone data [*Frith et al.*, 2004; *Stolarski and Frith*, 2006]; it is available from http://hyperion.gsfc.nasa.gov/Data_services/merged/. The dataset provides a nearly continuous time series of zonal and gridded (5° latitude by 10° longitude grid) monthly mean total ozone values between 60°S and 60°N (higher latitudes have data gaps during polar night) for the period from November 1978 to April 2009. Here we consider the period from January 1979 to December 2008 as well as the period from January 1979 to December 2005 for consistency with V07. Some data, particularly the data for August–September 1995 and May–June 1996, were missing. Zonal averages estimated from ground based total ozone measurements [*Fioletov et al.*, 2002] were used to fill the gaps.

3. Statistical Model

[6] A typical statistical model describing observations of monthly mean total ozone can be expressed in the form

$$\Omega(t) = a_0 + A(t) + Q(t) + S(t) + T(t) + X(t), \quad (1)$$

where $\Omega(t)$ denotes total ozone, t is the number of months after the initial time (taken here as January 1979), a_0 is the mean, $A(t)$ represents the seasonal cycle, $Q(t)$ the quasi-biennial oscillation (QBO), $S(t)$ the solar cycle, $T(t)$ the long-term trend, and $X(t)$ the residuals (noise). We use

$$A(t) = \sum_{j=1}^{4} a_{2j-1}\sin(2\pi jt/12) + a_{2j}\cos(2\pi jt/12), \quad (2)$$

$$Q(t) = [a_9 + a_{10}\sin(2\pi t/12) + a_{11}\cos(2\pi t/12)]w_{30}(t) + [a_{12} + a_{13}\sin(2\pi t/12) + a_{14}\cos(2\pi t/12)]w_{50}(t), \quad (3)$$

and

$$S(t) = [a_{15} + a_{16}\sin(2\pi t/12) + a_{17}\cos(2\pi t/12)]S_{107}(t), \quad (4)$$

where $w_{30}(t)$ and $w_{50}(t)$ are the equatorial zonally averaged zonal winds at 30 and 50 hPa respectively (http://www.cpc.ncep.noaa.gov/data/indices/), and $S_{107}(t)$ is the solar flux at 10.7 cm (http://www.drao-ofr.hia-iha.nrc-cnrc.gc.ca/icarus/). We use winds at both 30 and 50 hPa, because they are about 90 degrees out of phase, which allows a better representation of the QBO signal in total ozone. The $\sin(2\pi t/12)$ and $\cos(2\pi t/12)$ terms in $Q(t)$ and $S(t)$ represent the seasonal dependence of these effects. We do not include explicit terms in the statistical model for non-periodic variability such as that due to ENSO or annular mode variability, partly because we desire a parsimonious model and partly because those components are not predictable and so we prefer to consider them as part of our noise in our estimates of trend detectability as discussed by *Vyushin et al.* [2007].

[7] To describe the long-term trend in ozone that is related to ozone depleting species, the EESC is commonly used as a proxy in statistical models [e.g., *Dhomse et al.*, 2006; V07]. Long-term ozone changes can also be described using a piecewise linear trend (PWLT) with a turning point that is typically chosen in the second half of the 1990s [e.g., *Reinsel et al.*, 2002, 2005; *Miller et al.*, 2006]. The PWLT is closely related to the EESC approach since the EESC time series can be approximated to some extent by two linear functions, with the first slope equal to about 0.5 ppb/decade for the period before the EESC maximum in the second half of the 1990s and the second slope equal to about −0.2 ppb/decade for the period after the EESC maximum. (These numbers are for the lower stratosphere over midlatitudes [*Newman et al.*, 2007].) Therefore the EESC approach (at least over the 1979–2008 time period) is a particular case of the PWLT with a prescribed slope after the turning point equal to −40% of the slope before the turning point. Comparison of this prescribed slope to the slope estimated from the PWLT can be used as an indicator of how well the recent ozone changes agree with those expected from EESC. Therefore, we use either

$$T(t) = [a_{18} + a_{19}\sin(2\pi t/12) + a_{20}\cos(2\pi t/12)]EESC(t) \quad (5)$$

or

$$T(t) = a_{18}T_1(t) + a_{19}T_2(t), \quad (6)$$

where $T_1(t) = t$, for $0 < t \le n$, where n is time series length, and

$$T_2(t) = \begin{cases} 0, & 0 < t \le n_0, \\ t - n_0, & n_0 < t \le n. \end{cases} \quad (7)$$

[8] It should be mentioned that the estimates of V07 were based on an earlier version of the EESC [*World Meteorological Organization*, 2003] that is somewhat different from the more recent version [*Newman et al.*, 2007] used. The ratio between the EESC slopes after and before the turning point was −0.34 in that case (versus −0.4 used here), which affects the number of years required to detect the ozone trend expected from the EESC fit. In this study we use the new version of the EESC in Figures 1 and 2, the old

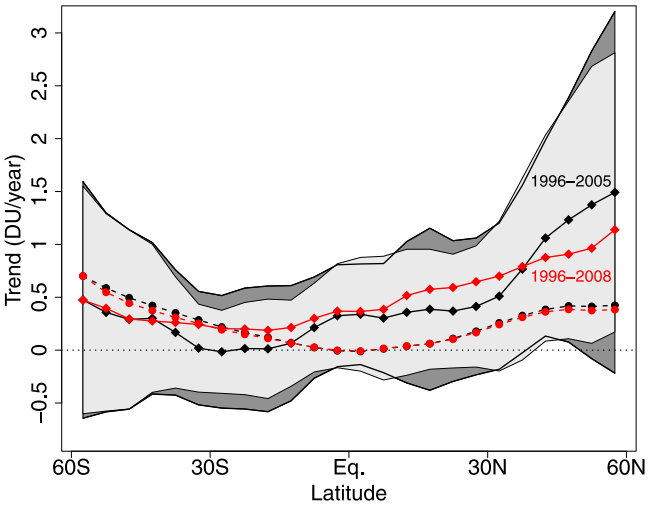

Figure 1. The EESC-based linear trend calculated for the declining part of EESC (solid black circles connected by the dashed line) is compared with the second (increasing) slope of the PWLT fit for the period 1996–2005 (black diamonds connected by the solid line), the latter with 95% confidence intervals calculated under the AR(1) (light grey region) and power law assumptions (dark grey region). The corresponding trend estimates for the period 1996–2008 are shown in red.

version in Figures 3, 4, and 5 for consistency with V07, and both versions in Figure 6.

[9] Figure 1 is an updated version of Figure 8 from V07. The circles show the total ozone trend estimate during the earliest stages of ozone recovery based on the EESC regression coefficient (i.e. coefficient a_{18} from equation (5)) and the diamonds show the growing part of the PWLT (i.e. $a_{18} + a_{19}$ from equation (6)). Trends for the period 1996–2005 (1996–2008) are plotted in black (red). The EESC based trend is more robust to the addition of the three extra years of data than is the PWLT because it is estimated using the full time series, and the extra three years are only about 10% of the previous 27 years. In contrast, the growing part of the PWLT is estimated using only the data after 1996, and the additional three years comprise 30% of the post-1996 data. Thus the PWLT estimate is more sensitive to interannual variability.

4. Temporal Aggregation Effect

[10] In this section we compare the relative validity of the AR(1) and power law statistical models for the total ozone residuals. Here we will address this question without actually fitting the statistical models to the power spectra.

[11] Our comparison exploits the distinctive behavior of the AR(1) and power law models under *temporal aggregation* as is done when, for example, creating an annual mean time series based on January to December averages of a monthly mean time series. We define the temporally aggregated time series

$$X^{(P)}(j) = \frac{1}{P}\sum_{t=1}^{P} X(t+(j-1)P),\ j = 1,2,\ldots,n/P,\ P \geq 1, \quad (8)$$

where $X(t)$, $t = 1, 2, \ldots, n$ is the original time series. In this notation, $X^{(12)}(1)$ would be the first value of an annual mean time series aggregated from the monthly mean time series $\{X(1), \ldots, X(12), X(13), \ldots\}$. *Vyushin* [2009] has shown that under temporal aggregation, for an AR(1) process with lag-one autocorrelation ϕ, the temporally aggregated process has lag-one autocorrelation

$$\text{AR(1)}: \quad \phi_{(P)} = \frac{\phi(1-\phi^P)^2}{P(1-\phi^2) - 2\phi(1-\phi^P)},\ 0 \leq \phi \leq 1. \quad (9)$$

In (9), $\phi_{(P)} = 0$ when $\phi = 0$, $\phi_{(P)} \to 1$ as $\phi \to 1^-$, and $\phi_{(P)} < \phi$ for $0 < \phi < 1$. The shape of $\phi_{(P)}$ as a function of ϕ is shown by the red curves in Figure 2. A similar approximate relation has been derived by *Kushnir et al.* [2006] for the comparison of monthly persistence of a monthly mean North Atlantic Oscillation (NAO) index with a prediction based on the AR(1) model fitted to the daily mean NAO index.

[12] By contrast, temporal aggregation has no impact on a power law stochastic process. More precisely, for a second order self-similar process, which can be regarded as the ultimate case of a power law stochastic process, we find [see, e.g., *Cox*,1984; *Taqqu*, 2002]

$$\text{Power law}: \quad \phi_{(P)} = \phi,\ 0 \leq \phi \leq 1. \quad (10)$$

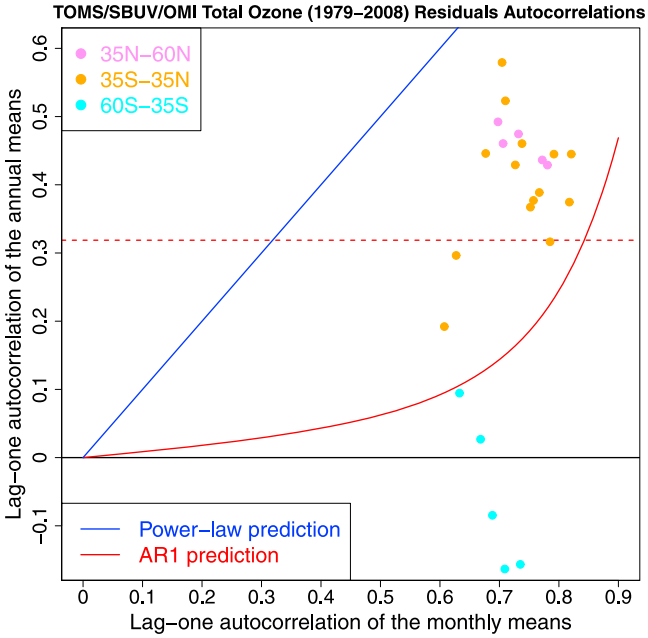

Figure 2. Scatterplot of $\phi_{(12)}$, which is the lag-one autocorrelation for the *annual* mean total ozone residuals, versus ϕ, which is the lag-one autocorrelation for the *monthly* mean total ozone residuals. The blue line represents $\phi_{(12)} = \phi$ from (10) for the power law statistical model; the red line represents $\phi_{(12)}$ as a function of ϕ from (9) for the AR(1) statistical model. The dots are color coded by region: "cyan": 60°S–35°S; "orange": 35°S–35°N; "violet": 35°S–60°N. Points above the horizontal dashed red line have significantly greater than zero lag-one autocorrelation for annual means at the 95% confidence level.

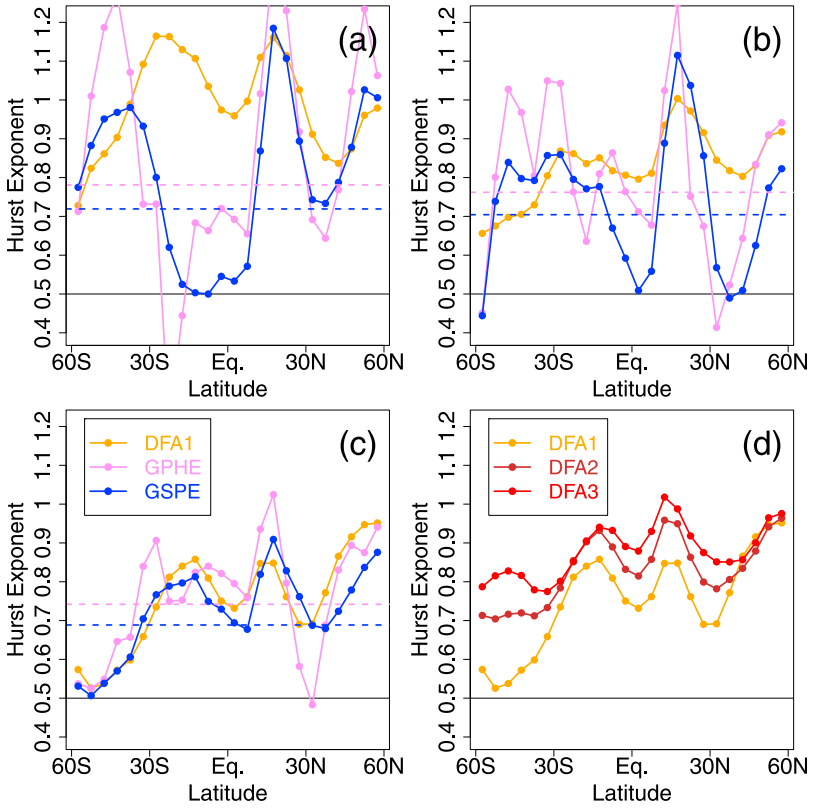

Figure 3. (a–c) Estimates of the Hurst exponent for zonally averaged TOMS/SBUV total ozone anomalies using DFA1 (orange curves), GPHE (violet curves) and GSPE (blue curves). (d) We also show the results of DFA2 (brown curve) and DFA3 (red curve). In Figure 3a only the annual cycle and the QBO have been filtered out, using linear regression. In all other panels annual cycle + QBO + solar flux + EESC have been filtered out. Timescales from 1 to 6 years have been used in Figures 3a and 3b and from 1 to 27 years in Figures 3c and 3d. The horizontal dashed lines in Figures 3a–3c show the upper 95% asymptotic confidence intervals for GPHE and GSPE for a Hurst exponent equal to 0.5, which corresponds to a stochastic process with white-noise-like low-frequency variability.

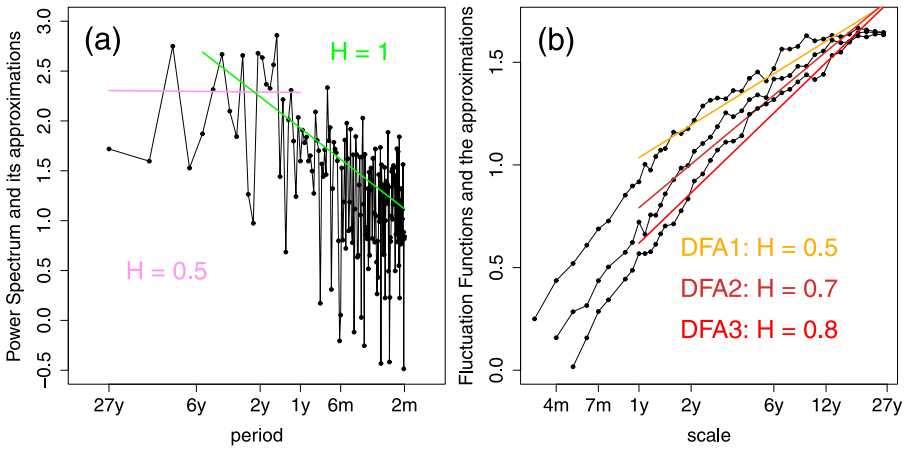

Figure 4. (a) Periodogram and (b) DFA1-3 fluctuation functions of the TOMS/SBUV total ozone anomalies obtained by filtering out annual cycle + QBO + solar flux + EESC from the 50°S–55°S zonal band. The periodogram, fluctuation functions, and their power law approximations are shown in log-log coordinates. The estimated values of H are rounded to one digit.

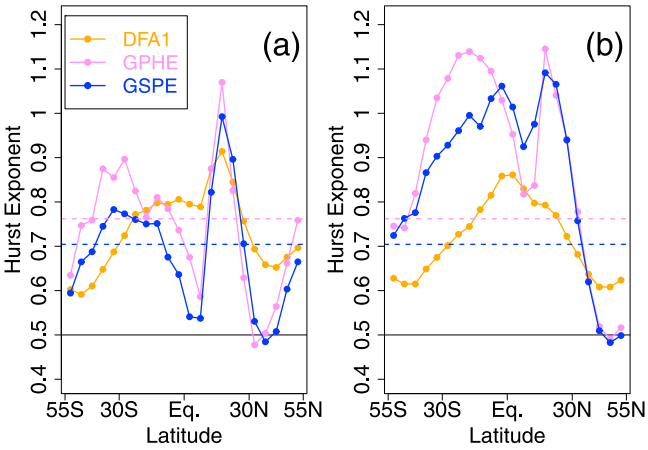

Figure 5. (a and b) Zonal averages of the Hurst exponent estimates obtained for the gridded TOMS/SBUV data using DFA1 (orange curves), GPHE (violet curves) and GSPE (blue curves) after filtering the annual cycle + QBO + solar flux + EESC and using timescales from 1 to 6 years. In Figure 5a the QBO has been filtered out by linear regression, whereas in Figure 5b it has been filtered out by the Wiener filter as in the work by K07.

This property, that the autocorrelation is independent of time aggregation, is certainly rather counterintuitive for climate processes.

[13] Equations (9) and (10) suggest a simple, and to our knowledge novel, test of the relative validity of the AR(1) and power law models: we examine the behavior of the lag-one autocorrelation under temporal aggregation in comparison with these equations. The results for the total ozone residuals are shown in Figure 2.

[14] Figure 2 is a scatterplot of the total ozone residuals' annual versus monthly autocorrelations ($\phi_{(12)}$ versus ϕ) for the period 1979–2008. Each point in this scatterplot represents a 5° wide zonal band between 60°S and 60°N with color coding for different regions (see the Figure 2 caption). The red and blue lines represent equations (9) and (10) respectively.

[15] All the points north of 35°S lie below the blue line and above the red line. Most of those points have annual lag-one autocorrelation significantly greater than zero. Therefore, for the total ozone residuals in the tropics and northern middle and high latitudes the AR(1) model (based on monthly mean ozone) provides a lower bound and the power law model an upper bound for persistence on monthly to inter-annual timescales. Interestingly, *Kushnir et al.* [2006] obtained a similar result for an NAO index on daily to monthly timescales, finding that *"the large monthly decorrelation values during winter are higher than expected from successive 30-day averages of a daily red-noise process."*

[16] The fact that the AR(1) and power law models provide lower and upper bounds, respectively, for the persistence of total ozone residuals on monthly to inter-annual timescales, and that the current standard practice in the ozone literature uses only the first of these bounds [e.g., *World Meteorological Organization*, 2007], motivates us to consider in detail the second bound. Fitting the power law model to time series residuals basically translates into the problem of Hurst exponent estimation, which is the main parameter of this model. *Toumi et al.* [2001], *Varotsos and Kirk-Davidoff* [2006], *Vyushin et al.* [2007], and *Kiss et al.* [2007] all used different spatial aggregation scales, different filtering techniques, different timescales, and different methods to estimate the Hurst exponent for total ozone residuals. In the next section we explain the differences between the two most extensive studies in this area [*Vyushin et al.*, 2007; *Kiss et al.*, 2007] and give a recipe for a robust Hurst exponent estimation procedure for geophysical time series.

[17] Physically, total ozone persistence stronger than can be explained by the AR(1) model might be attributed to low-frequency forcings such as the effect of SSTs on the tropical lower stratosphere [e.g., *Marsh and Garcia*, 2007] or low-frequency variability of planetary waves in northern midlatitudes [e.g., *Randel et al.*, 2002; *Hadjinicolaou et al.*, 2005]. The difference in the Hurst exponent values between the Northern and Southern Hemispheres can be explained by the difference in the main factors contributing to long-term variability there. In the Northern Hemisphere, dynamical factors play a substantial role in long-term ozone changes [e.g., *Randel et al.*, 2002; *Hadjinicolaou et al.*, 2005]. Dynamical factors contributed to low ozone values in the mid-1990s and to an accelerated ozone increase since then [e.g., *Harris et al.*, 2008]. We do not see similar long-term ozone variations in the Southern Hemisphere. However, a detailed attribution of total ozone persistence on

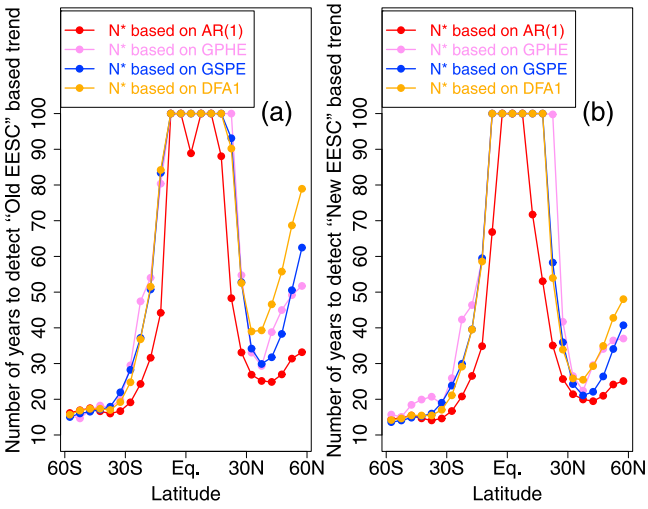

Figure 6. (a) The number of years since 2000 required to detect the old EESC based linear trend estimated for the declining part of EESC at the 95% significance level under the two alternative assumptions: AR(1) (red curve) and LRC (violet, blue, and orange curves based on GPHE, GSPE, and DFA1 respectively). The Hurst exponent estimation algorithm employed here is the same as in Figure 3c, but it was applied to the period 1979–2008. (b) The same as Figure 6a, but for the new EESC curve. Values higher than 100 years are plotted as 100 years. Note the similarity of the estimates in southern middle and high latitudes, but the large differences in the Northern Hemisphere.

various timescales requires detailed analysis and is beyond the scope of this study.

[18] Regarding the Hurst exponent estimates at different altitudes, we presume that the total ozone H estimate is dominated by the contribution from the region with the longest memory. For instance, the EESC seems to capture most of the long-term ozone fluctuations in the lower stratosphere except for the lowermost stratosphere (from the tropopause to about 18 km) in the Northern Hemisphere [*Yang et al.*, 2006]. Arguably, the contribution of that region to the total ozone in the Northern Hemisphere midlatitudes could be responsible for the super-AR(1) persistence of total ozone there. In contrast, long-term changes in the southern midlatitude lowermost stratosphere follow the EESC curve closely [*World Meteorological Organization*, 2007, Figure 3–11]. However, again these ideas warrant a more detailed analysis elsewhere.

5. Comparison of V07 and K07 Results

[19] There are seven differences between V07 and K07 in the way the Hurst exponents were estimated: (a) only TOMS data were utilized by K07, while V07 employed the merged TOMS/SBUV data set [*Frith et al.*, 2004]; (b) daily data was used by K07 versus monthly data used by V07; (c) DFA3 was used by K07 versus the Geweke-Porter Hudak estimator (GPHE [*Robinson*, 1995a]) and Gaussian semiparametric estimator (GSPE [*Robinson*, 1995b]) used by V07; (d) only the annual cycle and quasi-biennial oscillation (QBO) were filtered out by K07 versus annual cycle, QBO, solar cycle and the EESC trend by V07; (e) timescales from about a month to 6 years were used by K07 versus the frequency range corresponding to 1 to 27 years by V07; (f) V07 reported results for zonally averaged total ozone, while K07 reported the zonally averaged Hurst exponents for gridded total ozone; (g) the QBO was filtered by linear regression in the work by V07 and by the Wiener filter in the work by K07. We demonstrate here that the last four differences are key in explaining the different results reported.

[20] We first checked that the results for monthly TOMS/SBUV data are the same as for TOMS-only data. Therefore we can exclude item (a). Item (b) is irrelevant in these studies, because submonthly frequencies are not used. Providing a time series is generated by a power law stochastic process and is not contaminated by the presence of trends or periodicities, DFA and spectral methods (GPHE and GSPE) should give close estimates of the Hurst exponent [e.g., *Taqqu et al.*, 1995; *Vyushin and Kushner*, 2009]. However, climatic and meteorological time series do not usually have pure power law spectra. Often their spectra have different slopes for high and low frequency ranges [e.g., *Vyushin et al.*, 2007]. Moreover, they may contain periodic and quasiperiodic signals of various periods, as well as secular trends and data inhomogeneities caused for instance by changes in instrumentation, etc. In such circumstances an estimate of the Hurst exponent depends on the filtering applied to the time series and the choice of frequency (timescale) range [*Jánosi and Müller*, 2005; *Marković and Koch*, 2005; *Vyushin and Kushner*, 2009]. It was shown by *Vyushin and Kushner* [2009], using tropospheric and stratospheric air temperature as an example, that DFA and spectral methods give similar estimates provided equal frequency ranges — especially the lowest available frequencies — are chosen and that trends and periodic and quasiperiodic signals are filtered out. Below we show that these principles are also applicable to the analysis of total ozone.

[21] Figure 3 shows estimates of the Hurst exponent for the TOMS/SBUV zonally averaged total ozone anomalies estimated by DFA, GPHE, and GSPE for different combinations of filters and timescale ranges. Note that Figures 3–5 use the period 1979–2005 for consistency with V07. However, the corresponding figures (not shown) for the period 1979–2008 are very similar. We start by filtering the components that were filtered by K07, namely the annual cycle and the QBO. In Figure 3a the annual cycle and the QBO have been filtered out using linear regression on four annual cycle harmonics and the equatorial zonally averaged zonal winds at 30 and 50 hPa according to (2)–(3), and the Hurst exponents are estimated for timescales from 1 to 6 years (the intersection of the timescale ranges used by V07 and K07). The high-frequency cutoff was chosen by V07 equal to one year, because visual examination of the power spectra of total ozone residuals demonstrated that these spectra typically have two power law like regimes: intra-annual and inter-annual with a crossover around the annual timescale. The Hurst exponents estimated by DFA1, GPHE, and GSPE do not agree in this case. We then filter out the solar cycle using the solar flux at 10.7 cm and the EESC trend in addition to the annual cycle and the QBO (see Figure 3b). This brings the three curves closer together, but there are still noticeable differences between them. Comparison of Figures 3a and 3b reveals the effect of filtering of the solar cycle and the EESC trend, i.e. the effect of item (d).

[22] To fully comply with the recommendations of *Vyushin and Kushner* [2009] we then extended the timescale range up to 27 years thus including the lowest available frequencies. Most of the studies that have employed DFA have set the maximum timescale used to a quarter of the time series length. *Vyushin and Kushner* [2009] eliminated this limitation, demonstrating by means of Monte-Carlo simulations that the properties (bias and variance) of the Hurst exponent estimate obtained by DFA do not change appreciably when the longest available timescales are included. This actually makes DFA based estimates comparable to periodogram based results, for which the lowest available frequencies are typically included by default. The inclusion of the longest available timescales is also consistent with the fact that the Hurst exponent is defined only asymptotically [see, e.g., *Taqqu*, 2002]. Another motivation for the inclusion of the lowest available frequencies is that for the estimation of trend uncertainty (the trend is supposed to be filtered out prior to the Hurst exponent estimation) only the low-frequency behavior of the power spectrum matters [*Smith*, 1993]. Figure 3c shows that by including the lowest available frequencies we reach a noticeably closer agreement between the estimates provided by the different methods. The GPHE and GSPE Hurst exponent estimates (the violet and blue curves) in Figure 3c are the same as in Figure 4b from V07, which K07 used for comparison. Comparison of Figures 3b and 3c reveals the effect of item (e). The GSPE based estimates from Figure 3c were employed in the work by V07 for conservative trend uncertainty estimation.

[23] In the first three panels of Figure 3 we employed DFA of the first order (DFA1), because it can automatically filter out only discontinuities in time series, but not trends, and therefore is more similar to spectral methods than DFA of higher orders. K07 used DFA3, which automatically filters out local quadratic trends, and it is important to compare its results with the results of DFA1. We plot in Figure 3d the Hurst exponent estimates obtained by DFA1-3 after the annual cycle + QBO + solar flux + EESC have been filtered out and timescales from 1 to 27 years have been used. With the exception of the Southern Hemisphere middle and high latitudes the DFA results have qualitatively similar distributions with generally larger estimates obtained by higher orders of the DFA. We will explain the differences between the DFA results below in relation to Figure 4.

[24] Comparing Figures 3b and 3c one notices that the shift of the low frequency cutoff from 6 to 27 years has decreased the value of the Hurst exponent estimates over the SH middle and high subpolar latitudes. Let us now take a careful look at the power spectrum and DFA curves (fluctuation functions) of the total ozone anomalies for the 50°S–55°S zonal band. They are plotted in Figure 4. There are two scaling regimes in this power spectrum (see Figure 4a). The first is a high frequency one, which ranges from 2 months to somewhere between one and two years. The second, low frequency regime, ranges from about two years to the lowest frequency. If one fits a power law curve to the frequency range from 2 months to 6 years, which is the scaling range used by K07, then one obtains a Hurst exponent equal to about unity, as illustrated by the green line. In contrast, if a power law is fitted to the frequency range from 1 to 27 years, as illustrated by the violet line and as was done by V07, then the estimated Hurst exponent is about one half. This explains the difference seen at 50°S–55°S during the transition from Figure 3b to Figure 3c (the effect of item (e)). The sensitivity of the Hurst exponent estimates to the choice of the frequency range was stressed by V07 by contrasting their Figures 4 and 5.

[25] Figure 4b shows fluctuation functions and their best linear fits in log-log coordinates for DFA of the first, second, and third order. In agreement with Figure 3d the Hurst exponent estimate increases from 0.5 to 0.8 as one increases the DFA order from the first to the third. It is known that even for a power law stochastic process, DFA curves have two regimes: short timescale and long timescale [*Kantelhardt et al.*, 2001]. The Hurst exponent should be estimated by fitting a power law function to the long timescales. It is also known that the transition (crossover) point between the two regimes depends on the order of DFA: the higher the order, the larger the transition point [*Kantelhardt et al.*, 2001]. This phenomenon can be observed in Figure 4b. The transition point for DFA1 is located around a one year timescale, whereas for DFA3 it is close to two years. Thus using the same timescale range (from 1 to 27 years) for DFA1 and DFA3 the Hurst exponent is overestimated by the inclusion of the short timescale regime into the estimation domain for DFA3. Note, that although DFA fluctuation functions do not look exactly as straight lines in Figure 4b on timescales above one year, they are within confidence intervals of the best fitted straight lines. We conclude that K07 obtained significantly higher Hurst exponent estimates over the Southern Hemisphere high latitudes because they used a timescale range located in the high frequencies together with the third order of DFA. Both of these facts lead to an overestimation of the true Hurst exponent.

[26] The discussion above only partially explains the differences in the shape of the Hurst exponent distributions between V07 and K07. Figure 5 reveals the effect of items (f) and (g). Figure 5a shows the zonally averaged Hurst exponent obtained for the gridded TOMS/SBUV merged data set after the annual cycle + QBO + solar flux + EESC have been filtered out by linear regression. Timescales from 1 to 6 years have been used. Therefore Figure 5a is analogous to Figure 3b. One can notice that qualitatively the spatial distributions of the Hurst exponent are somewhat similar in these two panels. However, the zonally averaged Hurst exponents are generally lower than the Hurst exponents for zonally averaged ozone, in agreement with the theory of power law stochastic processes [*Granger*, 1980] and the results for the atmospheric general circulation [*Vyushin and Kushner*, 2009]. This phenomenon was also discussed by V07 with respect to their Figure 2. Nevertheless K07 compared in their Figure 7 V07's Hurst exponent estimates for the zonally averaged total ozone with their zonally averaged Hurst exponent estimates for the gridded data. When we recalculate the Hurst exponents plotted in Figure 5a for the timescale range from 1 to 27 years we get a picture very similar to Figure 3c with somewhat smaller values but with even better agreement between the methods (not shown).

[27] Figure 5b is analogous to Figure 5a, except that the QBO has been filtered out using the Wiener filter following K07. K07 linearly interpolated the total ozone anomaly power spectrum for frequencies in the range from 1.1 to 4.3 years. When we apply this filtering method to monthly ozone data the results are significantly affected, as can be seen by comparing Figures 5a and 5b. Remarkably, DFA1 and the spectral methods start to significantly disagree over several regions when the linear regression filtering of the QBO used by V07 has been replaced by the Wiener filter used by K07. All the methods demonstrate a relative boost of the Hurst exponent estimates over the tropics. Although K07 utilized daily data and a wider estimation frequency range, for which the 1.1 to 4.3 year range comprises a smaller portion, their results still could be affected by this rough filtering method.

6. Number of Years Required to Detect a Positive EESC Based Trend

[28] We now consider the impact of the different Hurst exponent estimates, and the revised EESC, on the number of years required to detect an EESC-attributable positive trend in ozone. The results are shown in Figure 6, expressed as the number of years after 2000 required to detect a trend at the 95% confidence level. As noted earlier, the LRC-based results can be considered as conservative upper limits on this date. The different LRC-based results all agree on the latitudinal dependence of the date. The use of the new EESC does not make much of a difference in the Southern Hemisphere, but in the Northern Hemisphere it reduces the number of years by between 5 and 30, depending on latitude. The updated results show that ozone recovery should be detectable by 2015 over 30°–60°S, and by between 2020

and 2025 over 30°–45°N, with the date increasing rapidly with latitude north of this range.

7. Summary

[29] Statistical analysis of total ozone satellite measurements is one of the major components of our knowledge about ozone variability and trends. Here we have considered the question of assigning proper confidence intervals to estimated total ozone trends in the past 30 years, which is a subject of both scientific interest and policy relevance. Until recently it has been standard practice to estimate confidence intervals of geophysical time series trends, including total ozone trends, using a first-order autoregressive model as a fit for the time series residuals. However, in recent years evidence has accumulated that the AR(1) model often underestimates the persistence of total ozone residuals. It follows that trend confidence intervals based on this model are also underestimated.

[30] A potential solution to this problem is to use an autoregressive model of order higher than unity. Such a model will keep the assumption of AR(1) that the power spectrum saturates to a constant in the low-frequency limit. However, *Vyushin et al.* [2007] demonstrated than even the AR(3) model underestimates the serial correlations of total ozone residuals in Northern Hemisphere high latitudes. Therefore even higher orders of the AR model would need to be used, which (a) raises the question of order selection, (b) causes a concern that the data might be overfitted, and (c) makes the problem of trend confidence interval estimation hardly tractable analytically.

[31] Alternatively, one can employ a model that (a) has the same number of parameters as AR(1), (b) is analytically tractable, and (c) is more conservative than any stationary autoregressive model. The power law model satisfies these conditions. However, one often cited issue with the power law model is the non-robustness of the power law exponent (Hurst exponent) estimates, which leads to different values of the exponent reported in different studies.

[32] In this study we have suggested that the power law model may be used as an upper bound for total ozone persistence on monthly to interannual timescales, and therefore as a basis for conservative (i.e. assuming strong natural variability) estimates of total ozone trend confidence intervals. We then explained the differences between the Hurst exponent estimates in the two most extensive studies, V07 and K07, and stressed the steps that need to be performed to obtain robust and sensible estimates.

[33] The major sources of the differences between the results of V07 and K07 for total ozone variability are the way the data were filtered and the frequency (timescale) range used to estimate the Hurst exponent. Secondary sources of differences were the estimators used and the use of gridded versus zonally averaged total ozone. V07 employed a low frequency range for two reasons. First, because only the low frequency variability affects the uncertainty of a trend, a proper estimation of which was the main goal of V07. Second, because the mathematical theory of power law stochastic processes is developed primarily for the asymptotic case, i.e. when the spectral density (autocorrelation function) scales by a power law for low frequencies (large time lags). K07 chose the intermediate range of the available timescales mainly because of conventional DFA requirements. This comparison shows that care is needed when estimation of the Hurst exponent is performed, and that none of the estimation methods should be used in isolation.

[34] Based on the analysis of the Hurst exponent, and using an updated EESC function, we predict that an increase in total ozone attributable to EESC should be detectable at the 95% confidence level by 2015 at the latest in southern midlatitudes, and by 2020–2025 at the latest over 30°–45°N, with the time to detection increasing rapidly with latitude north of this range.

[35] **Acknowledgments.** This research has been supported by the Canadian Foundation for Climate and Atmospheric Sciences and the Natural Sciences and Engineering Research Council. We thank the "R Foundation for Statistical Computing" for the *R environment* and David Pierce for the *ncdf* package. Dmitry Vyushin is grateful to Paul Kushner for fruitful discussions.

References

Cox, D. R. (1984), Long-range dependence: A review, in *Statistics: An Appraisal—Proceedings of a Conference Marking the 50th Anniversary of the Statistical Laboratory, Iowa State University*, edited by H. A. David and H. T. David, pp. 55–74, Iowa State Univ. Press, Ames.

Dhomse, S., M. Weber, I. Wohltmann, M. Rex, and J. P. Burrows (2006), On the possible causes of recent increases in Northern Hemispheric total ozone from a statistical analysis of satellite data from 1979 to 2003, *Atmos. Chem. Phys.*, *6*, 1165–1180.

Fioletov, V. E., G. E. Bodeker, A. J. Miller, R. D. McPeters, and R. Stolarski (2002), Global and zonal total ozone variations estimated from ground-based and satellite measurements: 1964–2000, *J. Geophys. Res.*, *107*(D22), 4647, doi:10.1029/2001JD001350.

Frith, S., R. Stolarski, and P. K. Bhartia (2004), Implication of version 8 TOMS and SBUV data for long-term trend analysis, in *Proceedings of the Quadrennial Ozone Symposium*, edited by C. Zerefos, pp. 65–66, Int. Ozone Comm., Athens, Greece.

Granger, C. W. J. (1980), Long memory relationships and the aggregation of dynamic models, *J. Econometrics*, *14*, 227–238.

Hadjinicolaou, P., J. A. Pyle, and N. R. P. Harris (2005), The recent turnaround in stratospheric ozone over northern middle latitudes: A dynamical modeling perspective, *Geophys. Res. Lett.*, *32*, L12821, doi:10.1029/2005GL022476.

Harris, N., R. Hudson, and C. Phillips (1998), SPARC/IOC/GAW assessment of trends in the vertical distribution of ozone, *SPARC Rep. 1*, 289 pp., World Meteorol. Org., Verriéres le Buisson, France.

Harris, N. R. P., et al. (2008), Ozone trends at northern mid- and high latitudes—A European perspective, *Ann. Geophys.*, *26*, 1207–1220.

Jánosi, I. M., and R. Müller (2005), Empirical mode decomposition and correlation properties of long daily ozone records, *Phys. Rev. E*, *71*, 056126, doi:10.1103/PhysRevE.71.056126.

Kantelhardt, J. W., E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde (2001), Detecting long-range correlations with detrended fluctuation analysis, *Physica A*, *295*, 441–454.

Kiss, P., R. Müller, and I. M. Jánosi (2007), Long-range correlations of extrapolar total ozone are determined by the global atmospheric circulation, *Nonlinear Processes Geophys.*, *14*, 435–442.

Kushnir, Y., W. A. Robinson, P. Chang, and A. W. Robertson (2006), The physical basis for predicting Atlantic sector seasonal-to-interannual climate variability, *J. Clim.*, *19*, 5949–5970, doi:10.1175/JCLI3943.1.

Marković, D., and M. Koch (2005), Sensitivity of Hurst parameter estimation to periodic signals in time series and filtering approaches, *Geophys. Res. Lett.*, *32*, L17401, doi:10.1029/2005GL024069.

Marsh, D. R., and R. R. Garcia (2007), Attribution of decadal variability in lower-stratospheric tropical ozone, *Geophys. Res. Lett.*, *34*, L21807, doi:10.1029/2007GL030935.

Miller, A. J., et al. (2006), Examination of ozonesonde data for trends and trend changes incorporating solar and Arctic oscillation signals, *J. Geophys. Res.*, *111*, D13305, doi:10.1029/2005JD006684.

Newman, P. A., J. S. Daniel, D. W. Waugh, and E. R. Nash (2007), A new formulation of equivalent effective stratospheric chlorine (EESC), *Atmos. Chem. Phys.*, *7*, 4537–4552.

Randel, W. J., F. Wu, and R. Stolarski (2002), Changes in column ozone correlated with the stratospheric EP flux, *J. Meteorol. Soc. Jpn.*, *80*, 849–862.

Reinsel, G. C., E. C. Weatherhead, G. C. Tiao, A. J. Miller, R. M. Nagatani, D. J. Wuebbles, and L. E. Flynn (2002), On detection of turnaround and recovery in trend for ozone, *J. Geophys. Res.*, *107*(D10), 4078, doi:10.1029/2001JD000500.

Reinsel, G. C., A. J. Miller, E. C. Weatherhead, L. E. Flynn, R. M. Nagatani, G. C. Tiao, and D. J. Wuebbles (2005), Trend analysis of total ozone data for turnaround and dynamical contributions, *J. Geophys. Res.*, *110*, D16306, doi:10.1029/2004JD004662.

Robinson, P. M. (1995a), Gaussian estimation of long range dependence, *Ann. Stat.*, *23*, 1630–1661.

Robinson, P. M. (1995b), Log-periodogram regression of time series with long range dependence, *Ann. Stat.*, *23*, 1048–1072.

Smith, R. L. (1993), Long-range dependence and global warming, in *Statistics for the Environment*, edited by V. Barnett and F. Turkman, pp. 141–161, John Wiley, Chichester, U. K.

Stolarski, R., and S. Frith (2006), Search for evidence of trend slowdown in the long-term TOMS/SBUV total ozone data record: The importance of instrument drift uncertainty, *Atmos. Chem. Phys.*, *6*, 4057–4065.

Taqqu, M. S. (2002), Fractional Brownian motion and long-range dependence, in *Theory and Applications of Long-Range Dependence*, edited by M. T. P. Doukhan and G. Oppenheim, pp. 5–38, Birkhäuser, Boston, Mass.

Taqqu, M. S., V. Teverovsky, and W. Willinger (1995), Estimators for long-range dependence: An empirical study, *Fractals*, *3*, 785–798.

Toumi, R., J. Syroka, C. Barnes, and P. Lewis (2001), Robust non-Gaussian statistics and long-range correlation of total ozone, *Atmos. Sci. Lett.*, *2*, 94–103, doi:10.1006/asle.2001.0042.

Varotsos, C., and D. Kirk-Davidoff (2006), Long-memory processes in ozone and temperature variations at the region 60°S–60°N, *Atmos. Chem. Phys.*, *6*, 4093–4100.

Vyushin, D. I. (2009), Statistical approximation of natural climate variability, Ph.D. thesis, Univ. of Toronto, Toronto, Ont., Canada.

Vyushin, D. I., and P. J. Kushner (2009), Power-law and long-memory characteristics of the atmospheric general circulation, *J. Clim.*, *22*, 2890–2904, doi:10.1175/2008JCLI2528.1.

Vyushin, D. I., V. E. Fioletov, and T. G. Shepherd (2007), Impact of long-range correlations on trend detection in total ozone, *J. Geophys. Res.*, *112*, D14307, doi:10.1029/2006JD008168.

Weatherhead, E. C., et al. (2000), Detecting the recovery of total column ozone, *J. Geophys. Res.*, *105*, 22,201–22,210.

World Meteorological Organization (2003), Scientific assessment of ozone depletion: 2002, *Rep. 47*, Global Ozone Res. Monit. Proj., World Meteorol. Org., Geneva, Switzerland.

World Meteorological Organization (2007), Scientific assessment of ozone depletion: 2006, *Rep. 50*, Global Ozone Res. Monit. Proj., World Meteorol. Org., Geneva, Switzerland.

Yang, E.-S., D. M. Cunnold, R. J. Salawitch, M. P. McCormick, J. Russell III, J. M. Zawodny, S. Oltmans, and M. J. Newchurch (2006), Attribution of recovery in lower-stratospheric ozone, *J. Geophys. Res.*, *111*, D17309, doi:10.1029/2005JD006371.

V. E. Fioletov, Environment Canada, 4905 Dufferin St., Toronto, ON M3H 5T4, Canada.

T. G. Shepherd and D. I. Vyushin, Department of Physics, University of Toronto, 60 St. George St., Toronto, ON M5S 1A7, Canada. (tgs@atmosp.physics.utoronto.ca; dmitry.vyushin@utoronto.ca)