



Trends in stratospheric ozone profiles using functional mixed models

A. Park¹, S. Guillas², and I. Petropavlovskikh³

¹Department of Statistical Science, University College London, London, UK

²Department of Statistical Science, University College London, London, UK

³CIRES, University of Colorado, Boulder, Colorado, USA

Correspondence to: A. Y. Park (ah.park.09@ucl.ac.uk)

Received: 17 April 2013 – Published in Atmos. Chem. Phys. Discuss.: 14 May 2013

Revised: 18 October 2013 – Accepted: 29 October 2013 – Published: 27 November 2013

Abstract. This paper is devoted to the modeling of altitude-dependent patterns of ozone variations over time. Umkehr ozone profiles (quarter of Umkehr layer) from 1978 to 2011 are investigated at two locations: Boulder (USA) and Arosa (Switzerland). The study consists of two statistical stages. First we approximate ozone profiles employing an appropriate basis. To capture primary modes of ozone variations without losing essential information, a functional principal component analysis is performed. It penalizes roughness of the function and smooths excessive variations in the shape of the ozone profiles. As a result, data-driven basis functions (empirical basis functions) are obtained. The coefficients (principal component scores) corresponding to the empirical basis functions represent dominant temporal evolution in the shape of ozone profiles. We use those time series coefficients in the second statistical step to reveal the important sources of the patterns and variations in the profiles. We estimate the effects of covariates – month, year (trend), quasi-biennial oscillation, the solar cycle, the Arctic oscillation, the El Niño/Southern Oscillation cycle and the Eliassen–Palm flux – on the principal component scores of ozone profiles using additive mixed effects models. The effects are represented as smooth functions and the smooth functions are estimated by penalized regression splines. We also impose a heteroscedastic error structure that reflects the observed seasonality in the errors. The more complex error structure enables us to provide more accurate estimates of influences and trends, together with enhanced uncertainty quantification. Also, we are able to capture fine variations in the time evolution of the profiles, such as the semi-annual oscillation. We conclude by showing the trends by altitude over Boulder and Arosa, as well as for total column ozone. There are great

variations in the trends across altitudes, which highlights the benefits of modeling ozone profiles.

1 Introduction

Trends in stratospheric ozone have been a concern for humans and the environment ever since the mechanism of ozone depletion was discovered (Crutzen, 1974; Molina and Rowland, 1974). As a result, the international community enforced the Montreal Protocol and its following amendments to curb emissions of ozone depleting substances (ODSs) (WMO, 2007, 2011). The discovery of the Antarctic ozone hole in the early 1980s (Farman et al., 1985; Solomon, 1999) was very recently followed by the discovery of a new ozone hole in the Arctic observed for an extended period of time (Manney et al., 2011). An increase in the occurrence of stratospheric ozone losses could dramatically increase human exposure to ultraviolet radiation, causing skin cancer and cataracts.

The link between ozone recovery and climate change also needs to be investigated. Indeed, there is new and stronger evidence for radiative and dynamical linkages between stratospheric change and specific changes in surface climate (WMO, 2011). In particular, Solomon et al. (2010) showed that stratospheric water vapor may have slowed the rate of warming by as much as 25%. Furthermore, part of the observed recovery in total ozone column levels may not be due to the Montreal Protocol restrictions on the production of chlorofluorocarbons (CFCs), but rather due to an increase in greenhouse gases (GHGs), which warm the troposphere, but increase stratospheric cooling that in turn may slow ozone depletion. Chemistry–climate models do not yet

simulate these interactions well, or do it with large uncertainties, and some joint effort by the CCMVal and CCMVal-2 projects focuses on intercomparisons of such models (Gillett et al., 2011). Having good estimates of trends from the lower to the upper stratosphere can potentially help disentangle this issue and improve numerical modeling.

We now discuss two ozone trend studies to clarify the relation of our work to these, and also to emphasize the possible improvements we make over them. Miller et al. (2006) analyzed profiles from 12 ozonesonde station located northward of 30° N. The data were collected from the 1970s until December 2003. For fractional Umkehr altitudes (quarter of Umkehr layer), the time series of ozone concentrations were regressed on monthly indices of the quasi-biennial oscillation (QBO), the solar cycle and Arctic oscillation (AO) as well as linear trend terms, with the use of an autoregressive noise. Miller et al. (2006) concluded that there was a change in the ozone trends around 1996, and that ozone in the lower stratosphere has been increasing from that approximate time. To borrow strength across a vertical profile, and thus improve trend estimation, Meiring (2007) was the first to analyze an entire set of ozone data measured at one location (Hohenpeissenberg) in a single model approach. Due to the irregular measurements in the ozonesonde data, Meiring (2007) initially interpolated the ozone data to a fine grid of regular intervals, followed by multivariate principal component analysis (PCA) of these vectors, and then used a cubic spline interpolation to retrieve continuous principal component functions. This provided a parsimonious representation of the major sources of ozone variations across altitudes. The coefficients of the leading principal components were used to investigate trends and the effects of QBO via a Smoothing Spline ANalysis Of Variance (SSANOVA) model (Gu, 2002). Even though Meiring (2007) mentioned the effects of the 11 yr solar cycle on the ozone levels, such a cycle was not directly used in the analysis. Instead, the evidence of the solar cycle was mentioned through the estimated time-dependent effect curves that exhibited peaks in 1970, 1981 and 1992, at the times the solar cycle was at its maximum. In Meiring (2007), the model was separately fitted for each month, therefore the QBO effects and the time trends were reported only for selected months, so borrowing of information across months was not possible. Finally, Meiring (2007) also mentioned the possible presence of more complex noise structures, but did not deal with it.

In this paper, we build a regression model that includes month, year, QBO, the solar cycle, AO, El Niño/Southern Oscillation (ENSO) cycle and Eliassen-Palm flux (EP flux) as additive terms. We use Additive Mixed effects Models (AMMs) to fit the covariate effects as nonlinear functions. The AMM is an extended version of the penalized regression spline, where each smooth term is represented by a linear combination of spline basis functions, and the coefficients vector is assumed to be random. The criterion we used for fitting is penalized least squares, which finds a compromise

between goodness-of-fit (fidelity to the data) and smoothing (avoiding roughness of the fitted functions). Smoothing parameters control the level of smoothing of the fitted functions. Data-driven smoothing parameter selection is known to be sensitive to misspecified error structure. Thus, we select the smoothing parameter as a variance component in the mixed effects model framework. As a result, we obtain more reliable estimates. Finally, we allow a more complex error structure that accommodates observed heteroscedasticity (here seasonal) in the residuals, since unexplained variations are present that are not purely noise.

Unlike Meiring (2007), we carry out an initial smoothing step via smoothing splines. We enhance the principal component decomposition by integrating two steps of data smoothing prior and after the PCA. Furthermore, by modeling all months in one regression setting, we are able to make use of information that is present across months. In this way the fitted curves of covariates are easier to interpret, as seasonal effects are already included in the analysis. By adding the covariates solar cycle, AO, ENSO and EP flux to the covariate QBO used in Meiring (2007), we are able to remove the effects of these influences in the ozone variations, and thus obtain trend estimates that correspond more genuinely to variations due to changing emissions of ODSs and GHGs.

The paper is structured as follows. In Sect. 2 we describe the Umkehr ozone data and represent them in terms of a basis function expansion in order to convert the discrete data into functional data. Section 3 explains the decomposition of these functional data using functional principal component analysis (FPCA) in order to perform dimension reduction. Sect. 4 focuses on finding the effect of proxies (QBO, solar, AO, ENSO and EP flux) on the modes of ozone variations via AMMs. In Sect. 5 we discuss the estimation results of our analysis and we derive trends in Sect. 6. Finally, Sect. 7 is devoted to conclusions and further discussion.

2 Data processing

2.1 Umkehr ozone data description

Umkehr daily ozone observations in Dobson units (DU) from January 1978 to December 2011 in Arosa and Boulder of latitudes 46°47'0" N and 40°0' 54" N, respectively, are used. The source of data is the WMO Ozone and UV Data Centre (WOUDC) and is publicly available at ftp://ftp.tor.ec.gc.ca/pub/woudc/Archive-NewFormat/UmkehrN14_2.0_1/. Ozone profiles are retrieved in sub-layers (where width is defined in a log pressure scale: a change in pressure between the top and bottom is a quarter of $\log(2)$ or approximately 1.2 km). Since the Umkehr method does not allow for independent information in high-resolution profiles, sub-layers are traditionally combined in thick Umkehr layers for further use in studies and archives (Petropavlovskikh et al., 2005). The layers are defined

according to the pressure level system. For example, the base pressure of layer 1 is approximately 0.0368 hPa, corresponding to 72.2 km of approximate height, while the bottom of layer 61 is at the sea-level pressure 1013.24 hPa. The total number of layers in the retrieved profile is 61. It fully covers the troposphere and the stratosphere and partially covers the mesosphere. If a station is located above sea level, e.g., Boulder, the information in the one or two bottom layers is not derived. Layers 1–28 (above 45 km) had no sensitivity to ozone variability due to limitations of the Umkehr method, so we discarded them from analysis. Hence, we focus on layers 29–60, corresponding roughly to altitudes ranging from 2 to 45 km.

At the beginning of the time series, the frequency of observations is considerably less than during the rest of the record, and the data are unequally spaced in time. Thus, we create monthly data by averaging out the daily record. Months for which no profile was observed created missing monthly profiles, e.g., for Boulder 11 monthly profiles were missing in the years 1978, 1979, 1982, 1983, 1998, 2003 and 2005, and for Arosa 4 monthly profiles were missing in the years 1978, 1986 and 2011. Finally, we removed the observations recorded over two volcanic periods: 1982–1983 (El Chinchón) and 1991–1993 (Pinatubo). Indeed, these observations were not corrected for aerosol interference, and therefore the profiles based on these two periods are erroneous.

2.2 Functional representation of ozone data

Even though the ozone profiles were divided into discrete layers, we view them as smooth curves, which reflect the degree of smoothing of the Umkehr method. With a functional representation that accommodates smoothness, the observations can be more realistically evaluated and understood, compared to a multivariate analysis that would not naturally account for such smooth dependence across altitudes. Thus, we achieve meaningful dimension reduction with our approach. For each station, functional ozone values corresponding to time i (each month and year combination) and layer j were observed:

$$y_{ij} \quad (i = 1, \dots, n; \quad j = 29, \dots, 60), \quad (1)$$

where y_{ij} is the ozone value, recorded at time i and layer j .

Let us briefly introduce how the functional ozone values are approximated by spline models. *Spline models* have become one of the most popular ways of fitting nonlinear functions. Splines are piecewise polynomials of degree r , with the polynomial pieces joining together at so-called *knots*, and they possess continuity conditions and a high degree of smoothness. Using a set of spline basis functions $\{\phi_k(x), k = 1, \dots, K\}$ and the associated coefficients $\{c_{ik}, k = 1, \dots, K\}$, we achieve smoothing as follows:

$$y_{ij} = y_i(x_j) + \epsilon_{ij}, \quad (2)$$

$$y_i(x) = \sum_{k=1}^K c_{ik} \phi_k(x),$$

where $y_i(x)$ is the smooth ozone profile of time i at altitude level x , and ϵ_{ij} are the associated space and time error. We choose a B-spline basis (de Boor, 2001) with the polynomial order of 4 (i.e., degree 3), so cubic B-splines are used as a basis function system. *B-splines* are known to provide a convenient basis for computational efficiency. Each B-spline basis function $\phi_k(x)$ is evaluated at knots and so we have to choose the number and locations of these knots to define the basis system. The degree of smoothing can be controlled by K , as small values of K can result in a smoother fit. Thus, we do not view a basis system as defined by a fixed number K of parameters, but rather we see K as itself a parameter that we choose according to the characteristics of the data.

The choice of the number (and locations) of knots is computationally expensive. In order to overcome the computational challenge, some authors proposed a roughness penalty approach, e.g., *smoothing splines* (Wahba, 1990) and *penalized regression splines* (Hastie and Tibshirani, 1990; Ruppert et al., 2003), which alleviates the heavy computational costs associated with knot selection. In the roughness penalty approach, the number K is chosen to be large enough to capture the maximum complexity of the function, but a penalty term involving a smoothing parameter takes care of excessive variations resulting from the large K . Roughness of the function $y(x)$ is often measured by the integrated squared second derivative of the function, i.e., $\int [D^2 y(x)]^2 dx$, where $D^m y(x)$ is the m th order derivative of the function $y(x)$. For each sample $y_i(s)$, we approximate the integral using a B-spline basis function expansion, i.e., $\sum_{k=1}^K \int [c_{ik} D^2 \phi_k(x)]^2 dx$.

According to de Boor (2001), the basis system that optimizes the least squares problem with a penalty term is the cubic B-splines system with knots placed at the observed data points. Thus, we placed knots at each layer excluding two end points consequently, in our situation $K = L + r - 1$ where L is the number of interior knots and r is the polynomial degree. We use the penalized least squares criterion to estimate the following coefficient vector for each station:

$$\begin{aligned} \min_{c_{ik}} & \sum_{j=1}^{32} \left(y_{ij} - \sum_{k=1}^K c_{ik} \phi_k(x_j) \right)^2 \\ & + \lambda \sum_{j=1}^{32} \sum_{k=1}^K (c_{ik} D^2 \phi_k(x_j))^2, \end{aligned} \quad (3)$$

where λ is a smoothing parameter. The first term quantifies goodness-of-fit (fidelity to the data), whereas the second term penalizes the roughness of the function (avoiding overfitting). Instead of the number K controlling the degree of

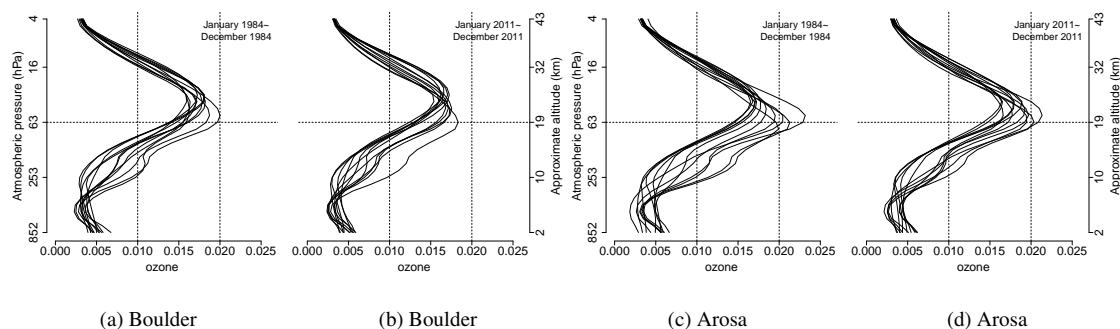


Fig. 1. Smoothed monthly ozone values (profiles) in DU (scaled by 1000) from two locations (Boulder and Arosa), based on the method of smoothing splines for the selected years of 1984 and 2011. The ozone profiles are represented as a function of altitude in km, or equivalently of atmospheric pressure in hPa. In each panel, vertical lines are drawn at 0.01 and 0.02. Horizontal lines are drawn at 63 hPa, which is approximately equivalent to 19 km in altitude.

smoothing, in the roughness penalty approach the smoothing parameter λ determines the level of smoothing. We used `smooth.basis` in the `fda` library in R to implement the estimation. Generalized cross validation (GCV) developed by Craven and Wahba (1979) helps us choose a smoothing parameter λ . The GCV scores were examined against a range of the parameter values. A plot of GCV scores against λ did not pin down a particular value for the parameter, as the scores were almost invariant regardless of the values of λ , provided that the values are approximately smaller than 10^{-5} . Thus we selected $\lambda = 10^{-5}$ for all samples of Boulder and Arosa.

The smooth monthly Umkehr ozone profiles as a function of altitudes for selected years (1984 and 2011) at two locations (Boulder and Arosa) based on the method of smoothing splines are displayed in Fig. 1. Following standard conventions in the atmospheric science community, we set the vertical axis as altitude and the horizontal axis as ozone values. We provide atmospheric pressure in hPa in addition to altitude in km, but the correspondence is approximate. The figure illustrates the functional nature of ozone variations (high correlation in ozone values along altitudes). As expected, the altitude-dependent ozone profiles exhibit substantial month-to-month variations.

3 Functional principal component analysis

The spline model approach presented in Sect. 2.2 achieves dimension reduction, but the dimension ($K = 32$) is still rather large. We now consider FPCA to reduce the dimension even further and represent each ozone profile in a more parsimonious way. FPCA has received a great deal of attention in the functional data analysis (FDA) literature as a device for dimension reduction, which is often an essential step for analyzing functional data. The Karhunen–Loève expansion, e.g., in Bosq (2000) tells us that

$$y_i(x) = \bar{y}(x) + \sum_{l=1}^{\infty} \xi_l(x) \theta_{il}, \quad (4)$$

where $\bar{y}(x)$ is the mean ozone profile, $\xi_l(x)$ is the principal component (PC) and θ_{il} is the uncorrelated random variable, referred to as the PC score, associated with i th sample and l th PC. Since the Umkehr ozone profiles were observed at discrete and dense layers (x_{29}, \dots, x_{60}), the PCs might be built as in the multivariate case using eigenvectors of the centered sample covariance matrix of the data and using an additional step of linear interpolation. However, this naive approach does not consider the functional nature of the data. The functional nature of the ozone data can imply that the PCs are smooth. We achieve smoothing of the PC by first smoothing the functional data by the method of smoothing splines as described in Sect. 2.2, and second projecting the PCs on a B-spline basis. Performing smoothing first also yields a smoothing of the covariance function of the data. However, if we use a large number for K , the level of smoothing is minimal. Silverman (1996) incorporated a penalty term into the orthonormality constraint imposed on the PCs to smooth the PCs even further. The detailed estimation procedure is presented in Appendix A1.

Using the truncated Karhunen–Loève expansion up to d (truncating the infinite expansion in Eq. (4) up to a finite number of components d), we achieve dimension reduction. We retain only the first $d = 5$ PCs, which are responsible for 99.6 % and 99.4 % of ozone variability at Boulder and Arosa. Typically, the PC gets rougher as the order l increases. The truncated Karhunen–Loève expansion is often understood as a way of eliminating noise, as higher-order PCs frequently represent the noise in the data. Also, PCs are often referred to as empirical basis functions because they provide basis functions to approximate the ozone profiles as seen from Eq. (4), and because they are driven by the data. With the empirical basis function $\xi_l(x)$, the PC score θ_{il} represents the relative

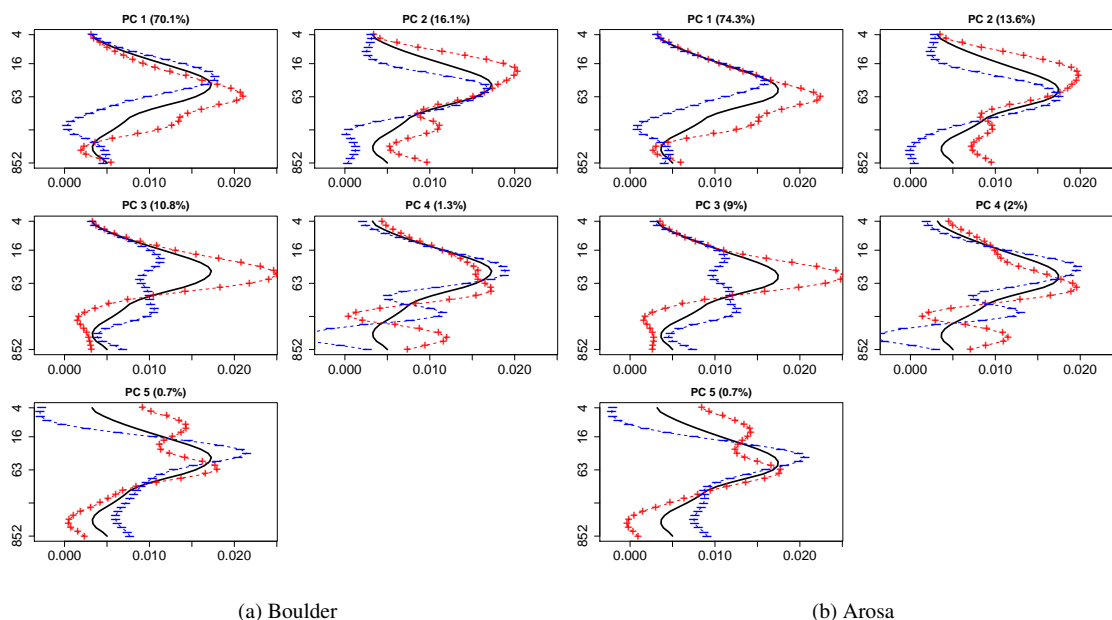


Fig. 2. The effects of five PC functions on the mean ozone profile. It displays the mean curve as a solid line, along with (+) and (−) indicating the exaggerated consequences of adding and subtracting a multiple of each PC. Dashed lines are drawn in addition to (+) to improve the visual quality. The x axis refers to ozone level in DU (scaled by 1000), and the y axis to atmospheric pressure (hPa). The variance contribution in % of each PC to total ozone variability is placed at the top of each panel.

contribution of the corresponding PC to ozone variations. Denoting the estimate of the PC from Silverman's approach by $\hat{\xi}_l(t)$, we compute the time series of the score vector θ_l for the l th PC for each station by

$$\theta_{li} = \int \hat{\xi}_l(x)(\hat{y}_i(x) - \bar{y}(x))dx \quad (i = 1, \dots, n), \quad (5)$$

where $\hat{y}_i(x)$ are the fitted Umkehr ozone profiles from the spline model in Eq. (2), and $\bar{y}(x)$ is the sample mean of the fitted ozone profiles. We have $n = 337$ for Boulder, as there were 11 missing months and 5 yr of volcanic periods were deleted ($337 = 408 - 60 - 11$), whereas we have $n = 344$ for Arosa, as there were 4 missing months and the volcanic periods were deleted ($344 = 408 - 60 - 4$). The time series of the vectors of the PC scores $\theta_1, \dots, \theta_5$ with $\theta_l = [\theta_{l1}, \dots, \theta_{ln}]$ are assumed to represent dominant temporal evolution in the shape of ozone profiles. Thus, they will be used in the subsequent statistical analysis, in which we aim to study the association between altitude-dependent ozone variability with specific time-dependent atmospheric forcings, such as QBO, solar cycle, AO, ENSO and EP flux (i.e., Sect. 4 includes regression on those external forcings).

The estimated five PCs and the associated PC scores are displayed in Fig. 3. If all PCs are found to be close to zero at a given altitude level, we can conclude that the profiles are close to its mean at this level, and relatively small variations are present. Figure 2 helps us inspect the effect of each PC on the mean ozone variation, since each PC represents variations

around the overall mean (we subtracted the overall mean profile before carrying out FPCA). The size of the perturbations around the mean curve, shown as (+)(−) in each panel, are computed by a multiple of each PC, i.e., $\bar{y}(x) \pm \delta \times \hat{\xi}_l(x)$. Conventionally, a standard deviation of each PC is widely used as the multiplier δ . However, we employ the same subjective choice of $\delta = 0.02$ here for all PCs to inflate the size of perturbations in order to enhance the visual quality.

It is useful to point out that the first five PCs are almost identical for Boulder and Arosa; see columns 1 and 3 of Fig. 3. Each PC shape is associated with sensitivity of the ozone profiles to major geophysical or chemical variations or combinations thereof. The contribution of these variations to ozone variability is not easy to disentangle. The analysis in the next sections will provide insight into these variations; for instance, PC 5 can be associated with the semi-annual oscillation. PC scores 3 and 4 for Arosa have larger variability than those PC scores for Boulder, which means that Boulder and Arosa have different dynamical contributions, despite the fact that the PC curves from the two different stations are almost identical. The first two PC scores show a clear annual cycle, which has been shown to be associated with both upwelling and in-mixing for the tropics, and thus by extension to the mid-latitudes through the Brewer–Dobson circulation (Konopka et al., 2010). Some outliers are detected in the time series of the PC scores. The beginning of the time series tends to have rather unstable measurements, possibly related to the fact that fewer measurements were available then.

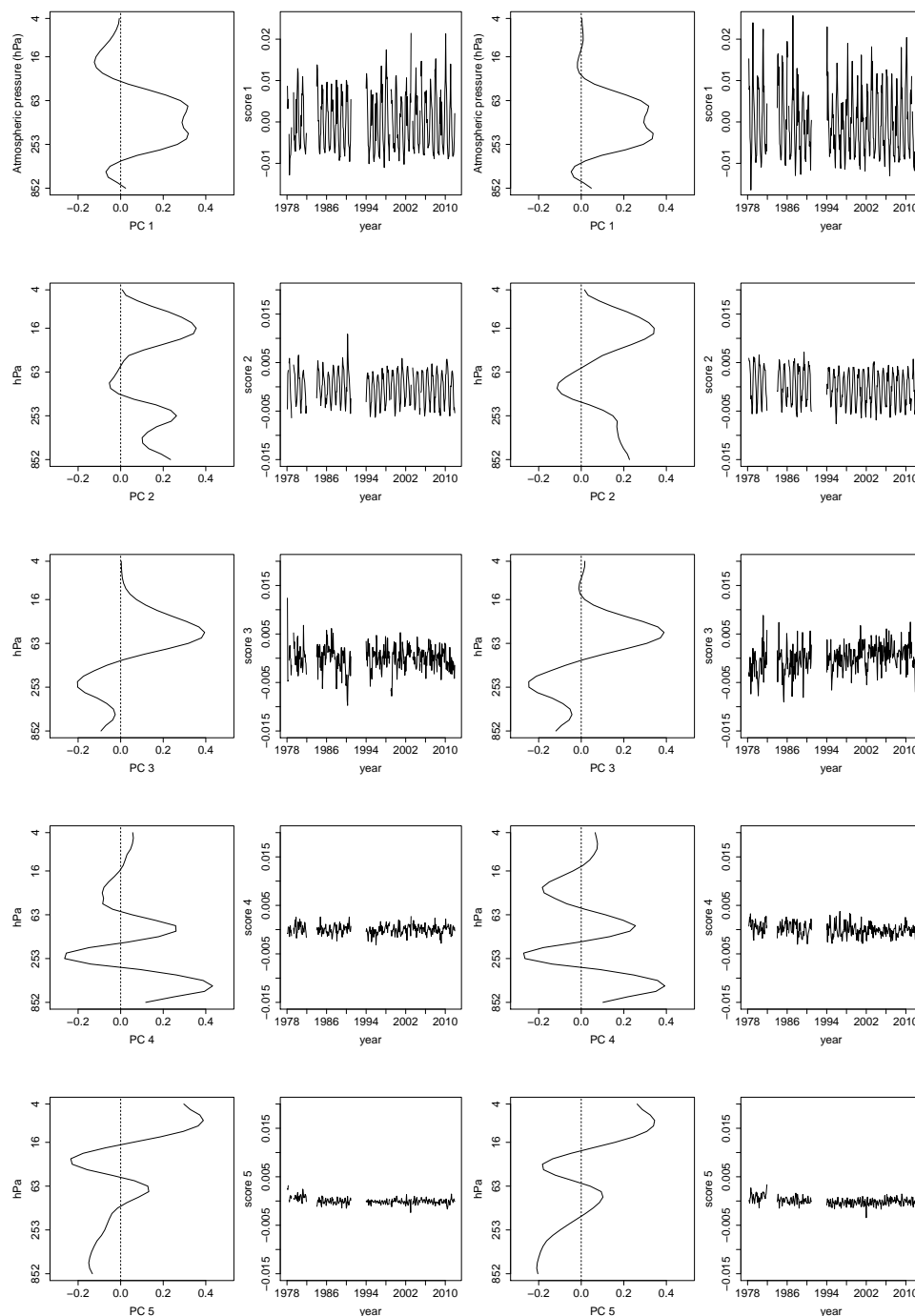


Fig. 3. The estimated five PC curves as a function of altitude in hPa, and the corresponding PC scores of monthly Umkehr ozone profiles from 1978 to 2011. First and third columns are the PC curves for Boulder and Arosa, respectively. Second and fourth columns represent the first five time series PC scores associated with the PC curve on their left-hand side. In the plots of the PC scores, two volcanic periods (1982–1983 and 1991–1993) are omitted.

When we use FPCA it is assumed that ozone values for each altitude level are normally distributed with a constant mean and variance. Histograms of the values for some altitude levels, e.g., layers 32, 33, 36, 50, 51, and 54, show

some skewness. However, the skewness is not present for layers where ozone is highly concentrated. Because we employ a functional approach, a transformation should be consistently applied to all altitude levels, and this is why we did

not consider the transformation. Since skewness occurs only at layers where ozone is not very high, we expect the bias in the FPCA step resulting from using non-normal data to be small.

We subtracted the overall mean ozone profile (altitude-dependent) prior to performing FPCA, thus a seasonal cycle is included in the covariance as well as in the time series of the PC scores. Indeed, variations in the PC scores are dominated by a strong seasonal cycle; see Figs. 6 and 7. This is why we include the month in the regression model in the next section (regression of the PC scores on the covariates), as the seasonal cycle in ozone is identified via a penalized regression cyclic cubic spline model. Further details of the results are provided in Sect. 5.

4 Modeling effect of covariates on ozone variations

In this section, we focus on finding the important sources of the unrevealed patterns and variations among the monthly Umkehr ozone profiles and on explaining these variations in terms of the relevant covariates. Here, we regress each PC score vector separately for each station on known external forcings. The main purpose of the regression is to partition each score vector into smooth components and random variability. The smooth components include month, year, two modes of variability of the QBO (QBO1 and QBO2), the solar cycle, AO, ENSO and EP flux. We use the method of penalized regression splines to fit each smooth component and select the smoothing parameters as a variance component in the mixed effects model framework. Section 4.1 introduces the covariates information, Sect. 4.2 focuses on introducing an additive model, where each smooth term is estimated by the penalized regression splines. Section 4.3 presents the relation between the penalized regression spline and the mixed effects model to show how smoothing is induced by the variance component. In Sect. 4.4, we carry out a variance function estimation to take into account the heterogeneous error structure.

4.1 Covariates

Here, we describe briefly the covariates (QBO, the solar cycle, AO, ENSO and EP flux) to be used in our analysis in Sect. 5. The QBO represents stratospheric zonal wind variations with a quasi-period of approximately 28 months. We use monthly QBO, available at <http://gcmd.nasa.gov/>. QBO was recorded at seven atmospheric pressure levels, 70, 50, 40, 30, 20, 15 and 10 hPa. We took the time lag in the QBO effect from the Equator to mid-latitudes into account by using a four-month lag; as a rule of thumb, 1 month per 10 degrees of latitude is often used in the literature. Furthermore, we used PCA to reduce the dimensionality (from 7 to 2) of the QBO records. Only the first and second dominant PC scores were kept. We denote the scores of the PC 1 as QBO1 and

the scores of the PC 2 as QBO2. PC 1 and 2 of the QBO data and their associated PC scores, QBO1 and QBO2, are presented in Fig. 5a. As seen from the shape of PC 1, QBO1 is mostly related to QBO winds at 30 hPa, whereas QBO2 is mostly related to QBO winds at 10 and 50 hPa.

Time series of other proxies are plotted in Fig. 5b. The solar cycle index represents the variations in the sun's activity, with an average period of about 11 yr. Daily solar cycle data (2800 MHz Series C) are available at <http://www.ngdc.noaa.gov>. Daily records were averaged out to create monthly solar flux. The AO index represents the major sea-level pressure variations North of 20° N of latitude. It does not show any particular periodicity. We use monthly AO data from <http://www.cpc.ncep.noaa.gov>. The ENSO index is associated with surface temperature and surface pressure variations over the tropical Pacific Ocean. We obtain monthly ENSO data from <http://www.esrl.noaa.gov/psd/enso/mei/table.html>. The EP Flux can be used as a proxy representing the planetary wave propagation to the upper stratosphere where it delivers the heat, and changes temperatures. The EP flux defines ozone transport from the Equator to high latitudes that builds up ozone in winter time, but then ozone experiences relaxation through photochemistry during the summer and early Fall. However, the rate of ozone destruction is fairly slow, therefore there is a correlation between ozone built up through March and the amounts of ozone observed in the following summer. This is why we used EP flux integrated from October to each consecutive month of the year; see Eq. (1) in Dhomse et al. (2003). We use NCEP EP flux re-analysis data (100 hPa, 45–75° N, monthly mean) available at http://www.awi.de/en/research/research_divisions/climate_science/atmospheric_circulations_old/projects/candidoz/ep_flux_data/.

4.2 Penalized regression spline

We aim to partition the PC scores for each station into the additive smooth components:

$$\begin{aligned} \theta_{li} = & c_l + g_{l1}(\text{Month}_i) + g_{l2}(\text{Year}_i) + g_{l3}(\text{QBO1}_i) \\ & + g_{l4}(\text{QBO2}_i) + g_{l5}(\text{solar cycle}_i) + g_{l6}(\text{AO}_i) \\ & + g_{l7}(\text{ENSO}_i) + g_{l8}(\text{EPflux}_i) + \epsilon_{li} \\ \epsilon_{li} \sim & N(0, \sigma_l^2) \quad (i = 1, \dots, n; l = 1, \dots, 5), \end{aligned} \quad (6)$$

where c_l is the overall mean and ϵ_{li} is the associated i.i.d error. $g_{lj}(F)$ is the smooth function of the covariate F_j and the l th PC score. We replace each smooth function g_{lj} with a linear combination of spline basis functions $\phi_{lk}(F_j)$ and the associated coefficients α_{ljk} :

$$g_{lj} = \sum_{k=1}^{K_j} \alpha_{ljk} \phi_{lk}(F_j), \quad (7)$$

and we estimate the coefficients $\alpha_{..k}$ for all l and j .

Following the spirit of a roughness penalty approach, the coefficients are estimated by minimizing the penalized least squares criterion with fixed smoothing parameters λ_{lj} :

$$\sum_{i=1}^n [\theta_{li} - c_l - \sum_{j=1}^8 g_{lj}(F_{ij})]^2 + \sum_{j=1}^8 \lambda_{lj} \int [D^2 g_{lj}(F_j)]^2 dF_j, \quad (8)$$

where the first term ensures the closeness of the estimate to the data, while the second term penalizes the curvature of each smooth function. For estimation, we use the *penalized regression cyclic cubic splines* for the month term and *penalized regression cubic splines* for the rest. The penalized regression spline is considered a generalization of the smoothing spline, with a more flexible choice of basis, penalties and knots. Unlike the smoothing spline, where knots are placed at each observation, in the penalized regression spline approach the number of knots is typically far less than the number of observations.

A splines basis system is determined by the amount and location of knots. However, in the penalized regression spline literature, it is known that knot selection does not have a large impact on the results of the model, if the coefficients are estimated by a balance between goodness-of-fit and the roughness of the function. Following Ruppert et al. (2003), we select $K_j = \min(n(F_j)/4, 40)$ and place the knots at fixed quantiles of the covariates, where K_j is the number of basis functions for covariate F_j and $n(F_j)$ is the number of the unique values of the covariate. Moreover, the choice of a class of a basis does seem to have very little impact on the fit of the model, provided that it has sufficient flexibility, numerical stability and appropriate mathematical properties. Thin plate penalized regression splines were also used to approximate the smooth functions g_{lj} in order to investigate sensitivity of the fit to the choice of the basis. However, the statistical results from the two models, involving two different basis systems, were almost indistinguishable. For a detailed fitting procedure of the penalized regression cubic splines for a given smoothing parameter, see Appendix A2.

4.3 Mixed effects model framework of penalized regression spline (AMMs)

The smoothing parameters λ_{lj} are unknown positive numbers, but they play a crucial role in fitting the penalized regression splines. Generalized cross-validation (GCV) is one of the widely used methods for selecting the smoothing parameter in the spline model literature. However, it is known that the smoothing parameters derived from GCV are heavily affected by a misspecified error structure, e.g., correlated errors (Krivobokova and Kauermann, 2007). The mixed effects model framework (Pinheiro and Bates, 2000) may help us achieve smoothing of the components in the regression model (6). Mixed effects models are an adaptation of regression models that incorporates a stochastic structure. In the mixed effects model framework of the penalized regression

splines, the model matrix \mathbf{B} in Eq. (A2), see Appendix, is partitioned into two parts (a fixed effects part denoted by \mathbf{B}_F and a random effects part denoted by \mathbf{B}_R), and the idea of penalization is incorporated into the covariance matrix of the random effects via a Bayesian approach (Ruppert et al., 2003; Wood, 2006).

Accordingly, we have the mixed effects model representation of the regression model in Eq. (6):

$$\begin{aligned} \boldsymbol{\theta}_l &= \mathbf{B}_F \mathbf{b}_l + \mathbf{B}_R \mathbf{u}_l + \boldsymbol{\epsilon}_l \\ \mathbf{u}_l &\sim N(0, \boldsymbol{\Omega}_{\lambda_l}) \\ \boldsymbol{\epsilon}_l &\sim N(0, \sigma_l^2 \mathbf{I}_n), \end{aligned} \quad (9)$$

where $\boldsymbol{\theta}_l = [\theta_{l1}, \dots, \theta_{ln}]^T$ is the l th PC score vector, and \mathbf{B}_F and \mathbf{B}_R are the model matrix corresponding to the fixed and random effects. \mathbf{b}_l is the coefficient vector corresponding to the fixed effects, e.g., coefficients of a constant and linear terms in the spline basis, and \mathbf{u}_l is the coefficient vector corresponding to the random effects. The fixed part is treated as unpenalized components and the random part is treated as penalized components; therefore, smoothing is induced by the covariance matrix of the random effects. We denote the covariance matrix of the random effects by $\boldsymbol{\Omega}_{\lambda_l}$ to emphasize that the matrix is determined by the vector of smoothing parameter $\boldsymbol{\lambda}_l = [\lambda_{l1}, \dots, \lambda_{l8}]^T$. Since \mathbf{u}_l is not fixed but assumed to be random, we predict it rather than estimate it. If we know $\boldsymbol{\Omega}_{\lambda_l}$ and σ_l^2 then we can predict \mathbf{u}_l using the conditional mean of \mathbf{u}_l given the response $\boldsymbol{\theta}_l$, i.e., $E(\mathbf{u}_l | \boldsymbol{\theta}_l)$. This conditional approach is known as the Best Linear Unbiased Prediction (BLUP); see, e.g., Robinson (1991). The detailed relation between the penalized regression splines and the mixed effects models is presented in Appendix A3. The estimation is implemented in `mgcv` with R-function `gam`.

Let us consider the notion of *effective degrees of freedom* for the regression model. The effective degrees of freedom (EDF) of each covariate assesses flexibility of the term in the regression model, and is closely related to the smoothing parameter. As the smoothing parameters increase from 0 to ∞ , the EDFs decrease smoothly from the maximum value (K_j defined in Sect. 4.2) to 1. If the smoothing parameter is large, then the model is less flexible, and so the fitted smooth curve has very few degrees of freedom. At the opposite extreme when the smoothing parameter is zero, then the penalty term in Eq. (8) vanishes; as a result, maximum of the EDF is achieved. When $\text{EDF} = 1$, the fitted curve is a straight line. Here, we cannot discriminate between the linear and the insignificant effects because the linear term is in the penalty null space, which means that the minimum value for the EDF is 1 for both the linear and insignificant effects. We employ a shrinkage method (Marra and Wood, 2011) as *variable selection* and it allows the discrimination. For variable selection we replace zero values in the penalty matrix \mathbf{S}_l in Eq. (A3) by a small value e . e was chosen to be very small so as not to affect the regression coefficients, except those in the penalty null space, e.g., constant and linear terms. As a

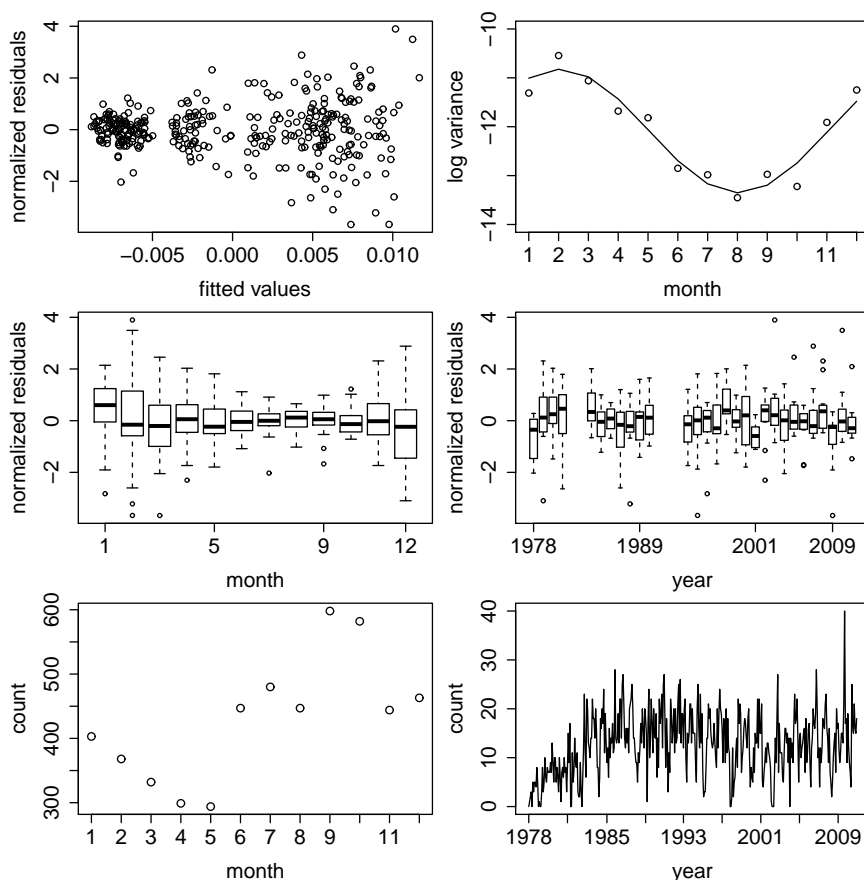


Fig. 4. Diagnostic plots of the regression model in Eq. (6) for PC score 1 of Boulder. Upper left panel: plot of normalized residuals versus fitted values. Upper right panel: log-variance of raw residuals computed for each month. The plots in the second row are box plots of normalized residuals grouped by month (left) and year (right). Bottom panels: counts of daily ozone observations per month (left) and year (right) as a time series.

result, regression coefficients are shrunk to zero if their associated smoothing parameter is large enough. In other words, when the EDF is less than 1, then we say the effect of the associated covariate is statistically insignificant. This approach achieves model selection without involving inference of the estimates. Marra and Wood (2011) present an extensive discussion about the variable selection for the penalized regression splines and provide guidance regarding its implementation for `mgcv` users.

4.4 Modeling heteroscedasticity

In the regression model (6), a particular form (i.i.d normal error) was assumed for the error term. Graphical and numerical summaries help analyze potential shortcomings of this assumption. Observations made in adjacent months might have stronger correlation than observations made in non-adjacent months. This potential correlation might not be completely captured by the time covariates, resulting in error correlation. In order to check the possible presence of serial correlation

in the error, residuals from the fit of the regression model (6) were graphically inspected. Autocorrelation plots of the residuals did not suggest that the error is correlated, so we do not consider autocorrelation here. Note that the original PC scores 1, 2 and 5 present strong autocorrelation, e.g., having approximate values of 0.7, 0.8 and 0.5 at lag 1 of Boulder. PC scores 3 and 4 have relatively weaker autocorrelations of 0.35 and 0.4, possibly due to the fact that PC 3 and PC 4 are not strongly related to time-dependent sources of ozone variation. Comparing the estimated autocorrelation of the residuals from the regression model and the autocorrelation of the original data, we are allowed to say that the penalized spline regression model eliminated autocorrelation in the original data.

Heteroscedasticity is present. Indeed, considerable variability remains even after the model is fitted, as seen in the plot of fitted values versus normalized residuals; see the upper left panel in Fig. 4. The plot reveals that the constant distributional assumption on the error is not appropriate. It gives

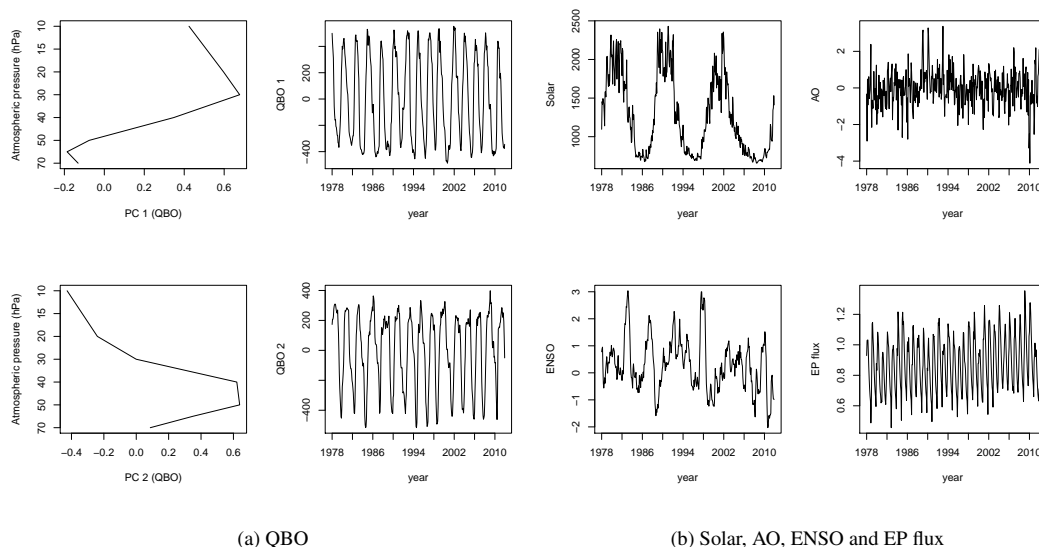


Fig. 5. Covariate: **(a)** QBO: PC 1 and PC 2 correspond to the first and the second dominant modes in the QBO; QBO1 and QBO2 are scores associated with PC 1 and PC 2. **(b)** Time series of covariates (solar, AO, ENSO and EP flux).

an indication that an extended model that includes a more complicated error system is needed to account for this remaining variability. As a first attempt to trace the cause of the observed variance pattern, the number of daily ozone observations used to create the average monthly data are counted. The counts are taken by month and year, and are displayed in the bottom panels of Fig. 4.

Log-variances of the normalized residuals (for PC score 1 of Boulder) are computed for each month and displayed in the upper right panel in Fig. 4. The plot shows high ozone variability in winter and spring months. The plot of the log-variances together with the box plot of normalized residuals grouped by month – see the plot in the second row and the first column of Fig. 4 – show that the residual variability has a strong annual cycle. It is possibly due to ozone transport in the upper and middle stratosphere associated with movement of jets (both in polar and subtropical) close and away from the station, and also due to the stratosphere–troposphere exchange. The errors in retrieved ozone in the upper stratosphere could also be related to the unaccounted stray light in the measurements that results in the underestimated values of retrieved ozone. The stray light in the band-pass is depleted more with increased total ozone, therefore the contribution of the out-of-band light becomes more significant; as a result, the errors can increase (Petropavlovskikh et al., 2011).

The log-variance versus month plot shows a periodic pattern. This pattern can be modeled by the sine and cosine waves:

$$\begin{aligned} \log(\text{Var}(\epsilon_{li})) &= 2\delta_{l1} \sin(2\pi \tilde{M}_i) + 2\delta_{l2} \cos(2\pi \tilde{M}_i) \quad (10) \\ \Leftrightarrow \text{Var}(\epsilon_{li}) &= \exp(2\delta_{l1} \sin(2\pi \tilde{M}_i)) \exp(2\delta_{l2} \cos(2\pi \tilde{M}_i)), \end{aligned}$$

where $\tilde{M}_i = \text{Month}_i/12$. The log transformation converts the multiplicative variance function into the additive one. We fol-

low Pinheiro and Bates (2000) and the convenient arbitrary choice of $2\delta_{l1}$ and $2\delta_{l2}$. To account for the heteroscedasticity, our model in (6) is replaced with a new model in which a complex error structure is assumed. In the new model, the error vector $\epsilon_l = [\epsilon_{l1}, \dots, \epsilon_{ln}]^T$ is assumed to be $\epsilon_l \sim N(\mathbf{0}, \sigma_l^2 \mathbf{\Lambda}_l)$. $\mathbf{\Lambda}_l \in \mathbb{R}^{n \times n}$ is a positive definite diagonal matrix with i th diagonal elements $\frac{1}{\sigma_l^2} \text{Var}(\epsilon_{li})$, where we define $\text{Var}(\epsilon_{li})$ in Eq. (10). By using the variance function in Eq. (10), we largely reduced the number of parameters to model heteroscedasticity. In other words, instead of estimating a whole set of diagonal elements in $\mathbf{\Lambda}_l$, we only estimate δ_{l1} and δ_{l2} .

5 Estimation results

In this section, we present the estimation results of the effects of the time covariates and the external forcings on the PC scores. Before presenting the results, we discuss two potential approaches that we considered to improve the model fit. Firstly, because we subtracted the overall mean ozone profile (altitude-dependent) prior to carrying out FPCA, a seasonal cycle is still included in the covariance as well as in the time series of the PC scores. We attempted to filter out the seasonal cycle in the FPCA step by subtracting the mean seasonal cycle from the profiles before performing FPCA. We computed sample mean ozone profiles for each month via the method of smoothing splines and used these sample means as estimates of seasonal and altitude-dependent mean profiles. When the seasonal cycle is filtered out, the PC scores no longer include the seasonal pattern, and month becomes statistically insignificant in the regression model for PC scores 1, 2, 4, and 5. Note that in the original analysis where the

Table 1. Estimated EDFs of each covariate on ozone scores. EDF indicates the influence and level of linearity of a covariate: EDF = 1 means linear relation; EDF = 0 implies non-significance; EDF > 1 means (possibly) nonlinear significance.

Scores	Mon	Year	QBO1	QBO2	solar	AO	ENSO	EPflux
Boulder								
score 1	7.5	0.5	0	0	1.2	0	1.7	1.6
score 2	8.2	1.2	0.7	0	1.9	0	0	1.7
score 3	6.6	3.4	2.8	0.9	1.5	0.7	1.3	2
score 4	4.9	1.8	1.9	0.9	2.3	0.9	0.8	0
score 5	7.9	4.5	1.6	0	1.4	0	0	1
Arosa								
score 1	6.9	1.4	0	0	0	1.2	0	1.7
score 2	7.9	2.2	0	1.2	1.1	0	0	0.9
score 3	4.2	1.9	0	0.2	0.9	0	0	0
score 4	4.5	1.9	2	0	0	1.2	0	2.7
score 5	7.5	4	0	0	0	0	0	1.5

seasonal pattern is not considered in the FPCA step (thus in the PC scores), month is statistically significant for all PC scores. In the end, however, the new statistical approach neither changes the fitted curves of covariates but month nor the estimated ozone trends. As a result, we decided to keep the seasonal cycle in the analysis, as it displays different features by score (e.g., semi-annual oscillation) that can be of scientific use. By removing seasonality in the first place, we would not be able to analyze such seasonal features (probably due to shrinkage).

Secondly, in order to investigate interaction effects between covariates (e.g., QBO–solar, AO–solar and QBO–AO), we used products of the values and created new variables, i.e., $QS = QBO1 \times solar$, $AS = AO \times solar$ and $QA = QBO1 \times AO$. Then, we fitted those three new variables (as interaction effects) via penalized regression cubic splines. For variable selection, following Marra and Wood (2011), we add a very small number to the penalty matrix affecting only the penalty null space, but not the penalty space. From the shrinkage approach, the EDF, an indicator of statistical significance, corresponding to the interaction terms, were all less than 1, with p-values larger than 0.2. Consequently, we conclude that the interaction effects are negligible.

5.1 Estimation results of covariate effects

Here, we report estimation results (numerical and graphical) of the covariate effects (from regression of the PC scores on the covariates). As a numerical result, we give the estimated EDFs of each covariate term; see Table 1. The EDFs indicate the level of nonlinearity, or equivalently the influence of the corresponding covariate on ozone (significance). When the EDF is 1, then the term is linearly related to the corresponding mode. If the EDF is smaller than 1, then the covariate is statistically insignificant according to our regression model. As expected, the geophysical covariates tend to be more lin-

early related (Miller et al., 2006) to ozone than months (annual cycle) or years (highlighting a trend change). Graphically, the covariate effects are represented as smooth curves (Fig. 6 and Fig. 7). The estimates of the covariate effects are shown as solid lines and their 95 % Bayesian confidence intervals are marked as shaded areas. Even though the penalty term in the fitting criterion compromises goodness-of-fit with the roughness of the curve, it biases estimates of parameters. As a result, confidence intervals based on a frequentist approach generally give poor results with regard to realized coverage probabilities. Alternatively, a Bayesian approach (Wahba, 1983; Silverman, 1985; Ruppert et al., 2003; Wood, 2006) among others, is widely used. Since the posterior distribution of the model parameters can be obtained, the construction of these confidence intervals is relatively straightforward. In addition to an easy implementation, simulations show that these Bayesian intervals have good observed frequentist coverage properties, resulting from the fact that they include both a bias and a variance component (Nychka, 1988; Wahba et al., 1995).

The confidence intervals marked as shaded areas should be interpreted with caution. Close to nominal coverage probabilities (here 95 %) are achievable for intervals obtained from a Bayesian approach when the interval performances are assessed across the function, provided that heavy over-smoothing is avoided (Marra and Wood, 2012). However, the intervals for smooth functions that are in the penalty null space, for instance straight lines, are problematic. When combined with the identifiability constraints necessary for the AMM estimation, estimates in the penalty null space often have confidence intervals that are almost of zero width at some points. Zero or narrow width implies that the estimation bias exceeds its variance. Consequently, it undermines the theoretical argument that a Bayesian type of interval achieves close to nominal coverage.

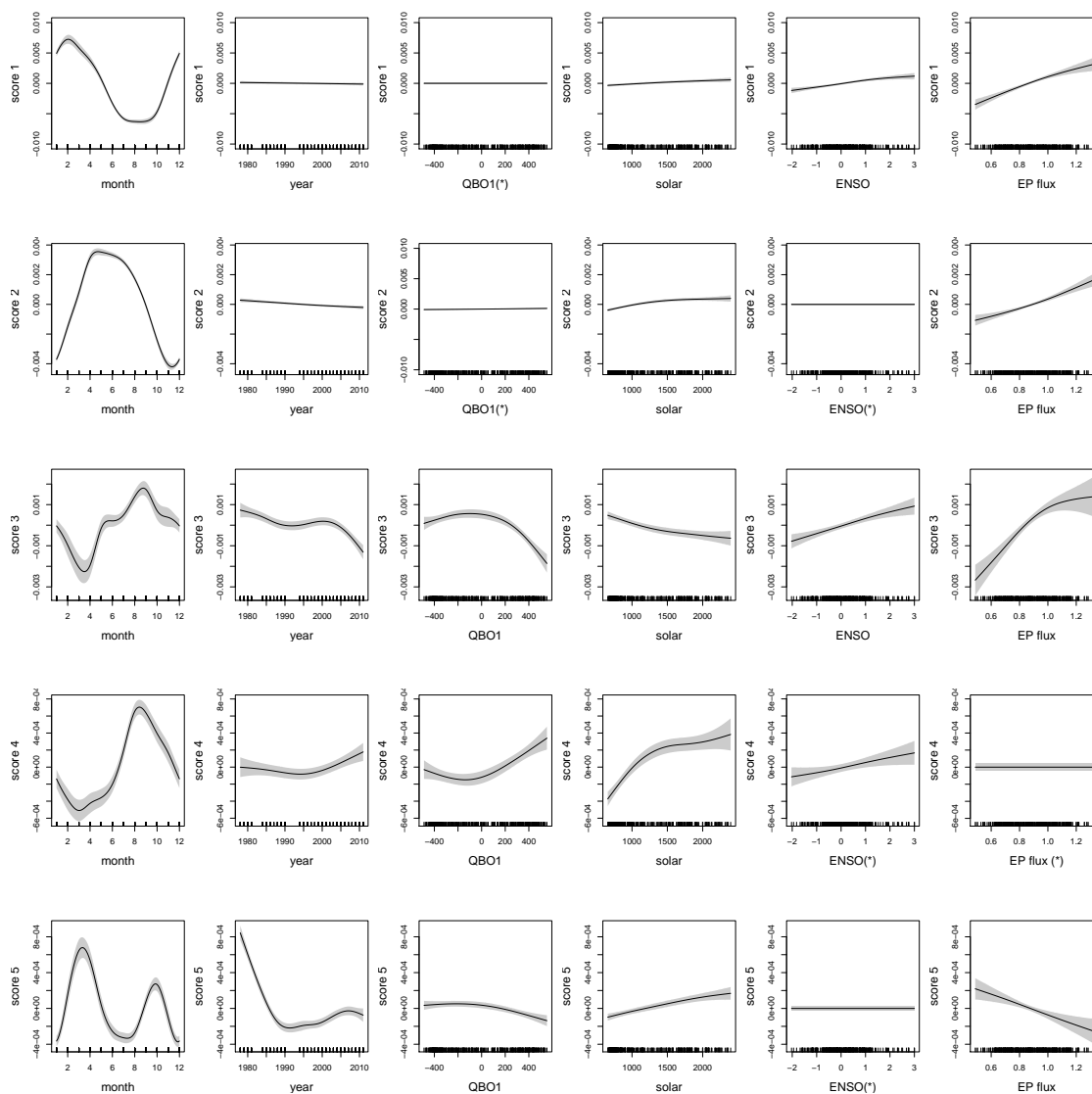


Fig. 6. Fitted smooth curves (black solid lines) and their 95 % Bayesian confidence intervals (shaded areas) from AMMs of Boulder for selected covariates (AO and QBO2 are excluded for all scores as none of these covariates are significant; see Table 1). The associated PC score and the covariate are marked as x and y labels. When the associated covariate is insignificant, under the 5 % significance level, (*) is added to the x label. The unticked years in the fitted curves of year represent volcanic periods omitted from our analysis.

Figures 6 and 7 show covariate fits into five PC scores for Boulder and Arosa, respectively. The covariates and PC scores are indicated by the labels of the x and y axes. The fitted month curve of PC score 5, both for Boulder and Arosa, shows very interesting features. Indeed, according to Garcia et al. (1997), the semiannual oscillation of stratospheric ozone has an M shape that peaks in March and October and mostly occurs in the upper stratosphere (score 5). To be able to pick up such a small variation in stratospheric ozone levels is an achievement of well-tuned AMMs. The fitted curves of year show the trends. Nevertheless, looking at these trends, we cannot claim here that a recovery in stratospheric ozone occurs for specific modes, but we have increased our confi-

dence that we can eventually pin them down with more time points (to improve estimates) when such changes become significant. We will give a detailed discussion of estimated ozone trends in Sect. 6.

It is also important to notice the difference between significant covariates that explain variation in PC scores for Boulder and Arosa. For example, QBO1 (corresponding to QBO winds at 30 hPa) is a prevalent covariate for PC scores 3, 4 and 5 for Boulder, while for Arosa, PC scores 2 and 4 are related to QBO2 (having maximum response at 10 and 50 hPa level) and QBO1, respectively. Also, one notices that PC score 3 (corresponding to ozone variation at altitudes between 63 and 16 hPa) for Arosa has no explanatory

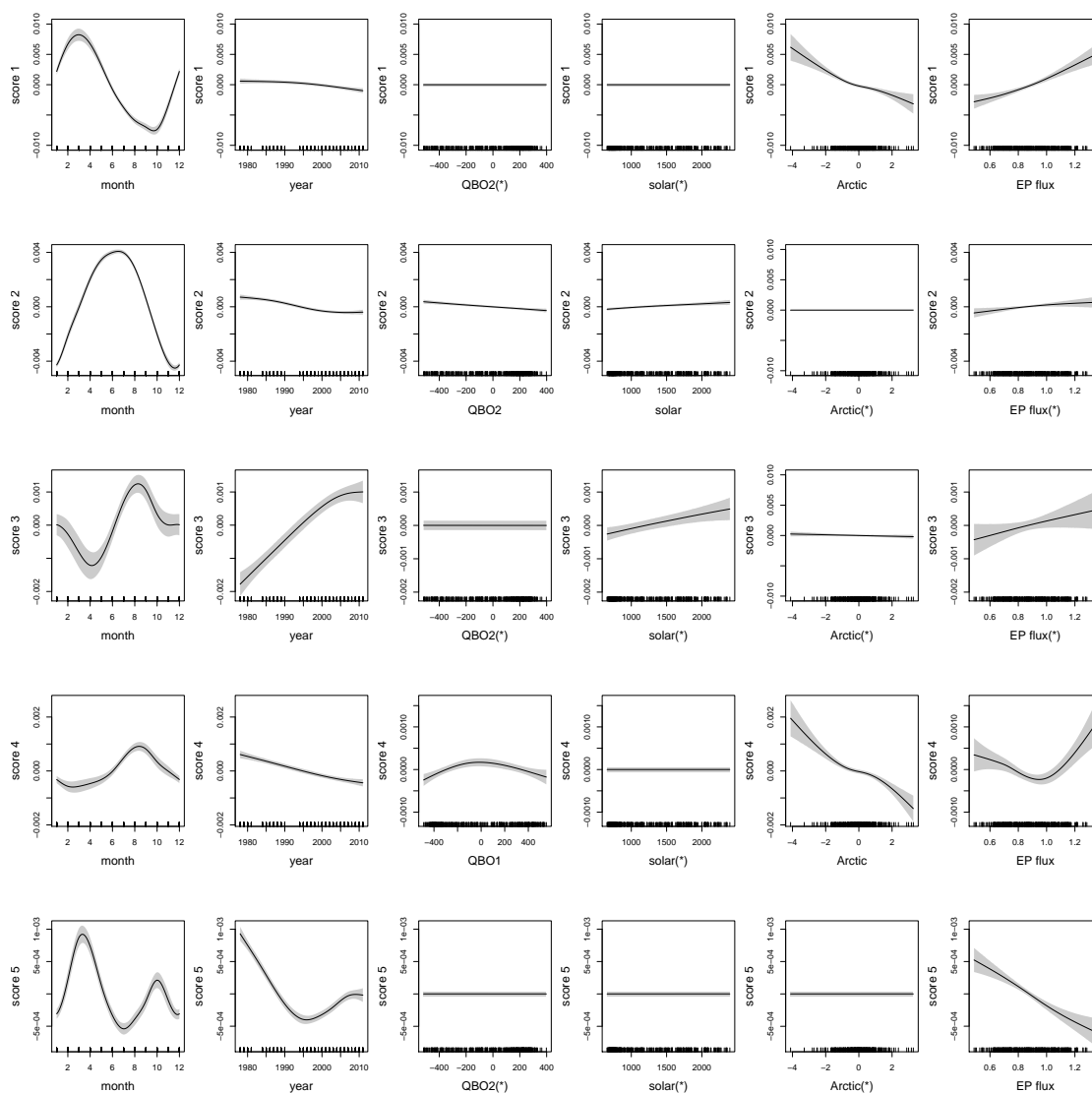


Fig. 7. Fitted smooth curves (black solid lines) and their 95 % Bayesian confidence intervals (shaded areas) from AMMs of Arosa for selected covariates (ENSO for all scores, and QBO2 for score 4, as well as QBO1 for scores 1, 2, 3 and 5, are excluded, as none of these covariates are significant; see Table 1). The associated PC score and the covariate are marked as x and y labels. When the associated covariate is insignificant under the 5 % significance level, (*) is added to the x label. The unticked years in the fitted curves of year represent volcanic periods omitted from our analysis.

parameters that are significant, except for the month and the year. On the other hand, PC score 3 for Boulder receives a significant contribution from additional covariates, such as QBO1, solar, ENSO and EP flux.

Another interesting difference is that PC scores for Boulder are associated with the ENSO covariate, while PC scores for Arosa show strong correlation with the AO, which probably makes sense since Arosa is farther north and thus has more influence from AO (Appenzeller et al., 2000). AO influences winter stratospheric ozone variability by changing the Brewer–Dobson stratospheric circulation, such that when the AO index is high, there is less wave breaking in the Po-

lar vortex; as a result, less ozone is transported to the high latitudes. The opposite happens when the AO index is low. Therefore, an increase in AO could result in apparent ozone depletion. However, some of the influence from the negative AO index is related to the blocking of the meteorological pressure systems (correlated with the Polar Eurasia index), and so creates very cold temperature over Eurasia (Steinbrecht et al., 2011). This can cause chemical ozone depletion and a change in ozone over Arosa associated with different meteorological patterns (up to 25–40 % springtime ozone decline) as compared to the patterns that govern the middle latitude ozone over the US (Boulder). There is also an

unresolved differentiation between NAO (North Atlantic oscillation) and AO indices, although many papers reference them interchangeably.

It is interesting to see that the solar covariate is significant only for PC score 2 for Arosa data. According to some authors (Reinsel, 2002; Newchurch et al., 2003; Dhomse et al., 2003; Hood and Soukharev, 2005; Miller et al., 2006), the solar cycle is usually one of the major contributors to ozone trend in the upper and middle stratosphere (PC scores 2, 3, 5). On the other hand, Frossard et al. (2013) showed that the contribution of the solar signal to total ozone is highly regional and seasonal. The reason is that the effect of the solar signal in the lower stratosphere is indirect, and of a mostly dynamical nature that corresponds to the changes in transport. However, in the upper stratosphere (or score 5), the contribution of the solar cycle should be closely related to the photochemical reactions that govern ozone concentrations. Further investigation is required to address this issue in the analysis of Arosa Umkehr time series. It is possible that the solar cycle is masked by other processes that have a significant contribution to ozone variability at Arosa station, or it is restricted by the sampling issues in Umkehr data that are limited by the fair-weather conditions. The peaks of the solar cycles are around 1981, 1992 (and 2003) and indeed match the two volcanic periods that we removed (1982–1983; 1991–1993). Nevertheless, we feel that this lack of information for large values of the solar proxy will have a limited impact on our analysis (probably making some confidence intervals wider).

We can compare our estimation results of covariate effects with other studies. First of all, we can compare them with Fig. 4 of Miller et al. (2006), which showed a negative influence of AO on ozone in the lower stratosphere, and we do also find a negative relationship for PC scores 1 and 4 of the Arosa data, associated with such altitudes. Furthermore, Miller et al. (2006) found a positive influence of solar on ozone in the upper stratosphere, and we do also find a slightly positive relationship for PC scores 2, 4 and 5 of Boulder and PC score 2 of Arosa, associated with such altitudes.

Secondly, the sensitivity of total ozone variability to the QBO signal has been discussed in Frossard et al. (2013). They found that QBO 30 hPa (QBO1) has a negative contribution to total ozone variability over the US (Boulder) and some of Europe (Arosa), and the contribution is significant (smaller p value than 1 %). Although the results are obtained for total ozone time series we can expect similar results for Umkehr ozone profiles at altitudes of ozone maximum (Umkehr layers 4 and 5, or between 64 and 16 hPa, or PC score 3). From our AMM analysis of PC score 3 for Boulder, we find that negative values (reduced ozone) correspond to the positive range of the QBO1 covariate (although mostly nonlinear). However, for Arosa's PC score 3 the QBO1 covariate has EDF of 0.2 (Table 1), thus it is insignificant. In addition, Frossard et al. (2013) found that the total ozone exhibits an increase due to variability in QBO 50 hPa (QBO2)

over the US, while a change in total ozone is found to be negative over Europe. However, results might be less significant; see Fig. 6 in Frossard et al. (2013). Our AMM analysis of PC score 3 for the Boulder and Arosa data does not find a significant contribution from QBO2 (EDF = 0.9 and 0 respectively). However, for PC score 2 the QBO2 covariate shows a significant contribution (EDF = 1.2) for Arosa but no contribution (EDF = 0) for Boulder.

Thirdly, we compare our analysis with the analysis of total ozone data in Rieder et al. (2010) and Rieder et al. (2013). Our AMM analysis finds a significant contribution of ENSO to Boulder (US) ozone variability for PC score 3 (where ozone is maximum), which is similar to the result of Rieder et al. (2010) analysis that also found a positive contribution over the US region. Note that the contribution is significant only for Winter and Spring months. According to Rieder et al. (2013), over Europe the total ozone analysis showed a negative effect of ENSO, but not significant, which is supported by our analysis of PC score 3 for Arosa that shows the non-significant influence. Rieder et al. (2013) also indicated the seasonal difference in the ENSO contribution. For instance, in Winter months (see their Fig. 2) contributions to total ozone are similar to the case for the full year, but only for Europe the p value is very low (meaning significance). For Spring (see their Fig. 3), both the US and Europe would gain a positive contribution, but only the US would have a small p value (significant). Since our analysis does not assess the ENSO contribution by season, we cannot fully compare our yearly ENSO results to the Rieder et al. (2013) findings. Rieder et al. (2013) also gave an analysis of the NAO signal contribution to the total ozone column data set (based on the SBUV satellite time series) and found that negative ozone changes are related to NAO over the US and European region, and that the changes are significant in both areas; see Fig. 8 in Rieder et al. (2013). Since our analysis uses AO, there may be a difference between AO and NAO patterns.

The model that includes a more complex stochastic structure (heteroscedastic error) led to modest (not that significant) changes in the shape of estimated curves. Nevertheless, Fig. 10a and c support the need for the more complex model. 95 % Bayesian confidence intervals in Fig. 10c take the dissimilar variability among months into account, which manifests itself in a greater uncertainty of the estimated curve through February to April but a decrease through Autumn. Thus, a periodic pattern in the log-variances of residuals is well adapted to the measure of the uncertainty. The parameters in Table 2 show the magnitude of seasonal pattern in the variance of the residuals, or equivalently, the extent of heteroscedasticity in the residuals. Besides, the fact that the AMMs treat the noisy part of a curve as random might lead to more robust estimates, reducing bias and avoiding overfitting. Note that the fitted curves of the covariate effects are smooth. It might be more realistic to assume that the curves of the covariate effects are smooth.

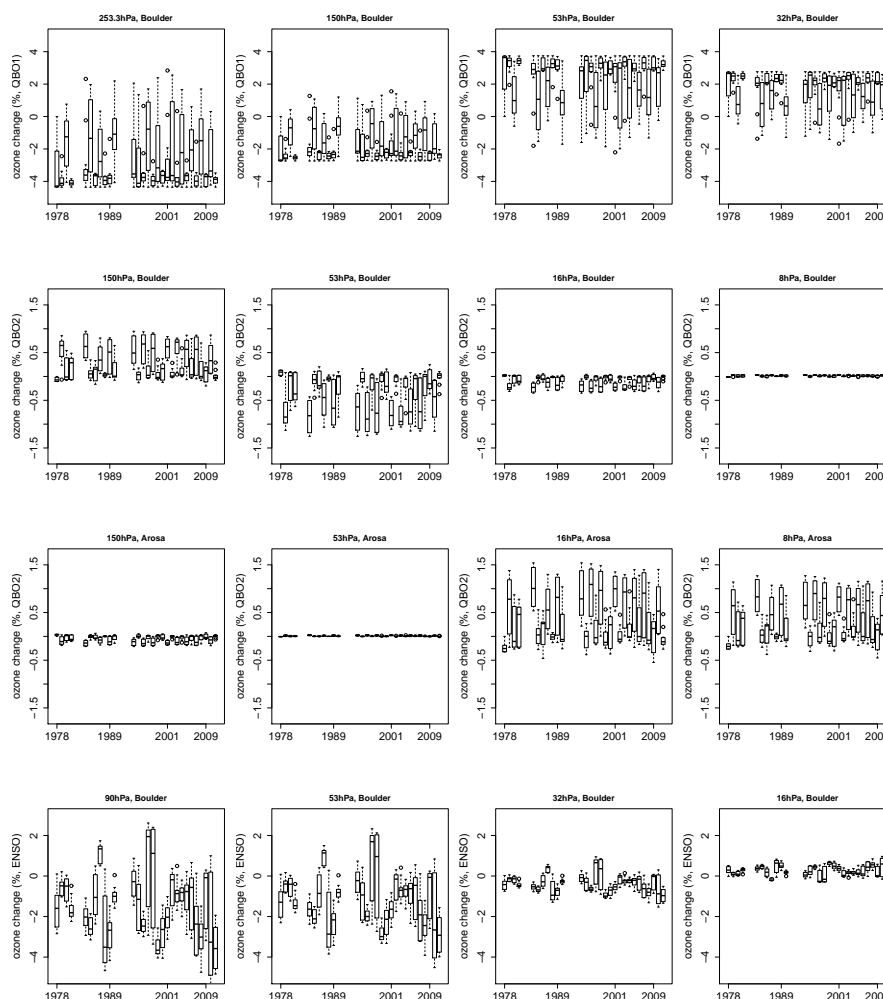


Fig. 8. Monthly contributions of covariates QBO1, QBO2 and ENSO for Boulder, QBO2 and ENSO for Arosa (covariates indicated in y labels) to changes in ozone profiles for selected altitudes (indicated on top of each plot with station name). The contributions (%) are grouped by year and displayed as box plots.

Table 2. Estimates of variance component parameters δ_1 and δ_2 in Eq. (10). The parameters shows the magnitude of seasonal pattern in the variance of the residuals.

Scores	Boulder		Arosa	
	δ_1	δ_2	δ_1	δ_2
score 1	0.48	0.48	0.45	0.35
score 2	0.32	0.19	0.20	0.30
score 3	0.44	0.36	0.28	0.32
score 4	0.13	0.26	0.11	0.24
score 5	0.45	0.13	0.13	0.093

Explained deviances D_l in Table 3 demonstrate the quality of the regression model fit and are defined as

$$D_l = 1 - \frac{\sum \tilde{\epsilon}_{li}^2 / (n - df_l)}{\sum (\theta_{li} - \bar{\theta}_{li})^2 / (n - 1)}, \quad (11)$$

where $\tilde{\epsilon}_{li}$ is the normalized residuals defined in Eq. (12) and df_l is the sum of the EDFs of all smooth terms for the l th PC score. The explained deviance quantifies the portion of the variations that can be explained by our model. These numbers are high for scores 1, 2 and 5, but low for scores 3 and 4. This may be due to the nature of the variations associated with PC 3 and PC 4. These components may be associated with short-term dynamics that are neither easily captured by seasonal changes, significant trends, nor by the main medium-term variations of the encapsulated atmosphere considered here. Changes in dynamical parameters would come from radiative and chemical mechanisms

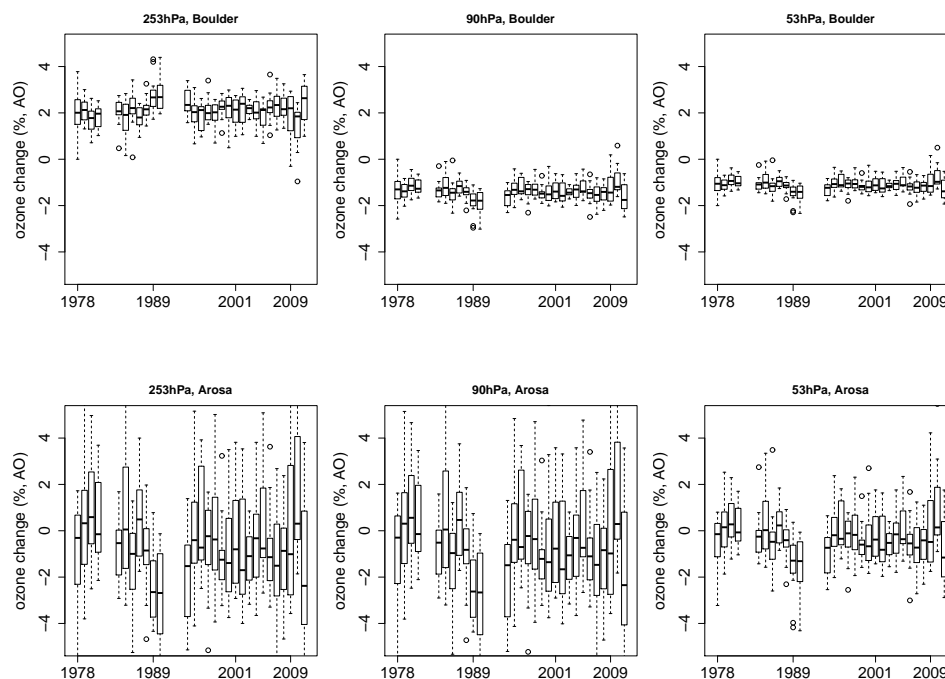


Fig. 9. Monthly contribution of AO to changes in ozone profiles for selected altitudes (indicated on top of each plot with station). The contribution (%) is obtained as monthly data, but we group them by years to remove seasonal variability. The contributions (%) are grouped by year and displayed as box plots.

Table 3. Explained deviance D_l (%) in Eq. (11) and the estimates of the noise levels σ_l in Eq. (9). Explained deviance indicates overall quality of model fit.

Scores	Boulder		Arosa	
	D_l	σ_l^2	D_l	σ_l^2
score 1	94	5.88×10^{-6}	88	1.11×10^{-5}
score 2	91	1.04×10^{-6}	91	1.34×10^{-6}
score 3	40	3.89×10^{-6}	31	5.82×10^{-6}
score 4	39	6.09×10^{-6}	44	9.77×10^{-7}
score 5	67	1.62×10^{-6}	50	3.19×10^{-7}

imposed by increases in GHGs. Increases in GHGs can warm the stratosphere, and thus directly affect ozone destruction rates that are temperature-dependent (so-called super-recovery). It could also change the transport patterns geographically and seasonally that can alter ozone at the lower stratosphere where dynamics play an important role. Therefore, it is important to study these contributions and estimate their contribution to ozone variability to improve the regression fit.

Figures 11 and 13 show Umkehr monthly ozone profiles and the estimated monthly profiles using the fitted PC scores from the AMM (under the heteroscedastic error assumption) for selected altitude levels. Note that the estimated monthly

ozone profiles at an altitude of x are computed by Eq. (4), but truncating the infinite series at $l = 5$ and $\hat{\theta}_{li}$ replacing θ_{li} , where $\hat{\theta}_{li}$ are the fitted PC scores from the AMMs. The fit is good, as variations are captured by our explanatory variables. Even though model flexibility of the AMMs allows complicated fitting of time series, it does not fully capture high and sharp peaks. We believe that the random spikes not accounted for by the regression model correspond to a smaller scale variability that is not captured by proxies, and they can be related to sudden stratospheric warming (Sofieva et al., 2011), wave breaking (Holton, 1983) and other solar high proton events (Seppala et al., 2004) that are not included in the model. There is some information in the upper stratosphere (e.g., 3.9 hPa, Arosa) that is not captured by the regression model. In order to improve the fit, it is important to study dynamical processes to find their effects on long-term changes in ozone distribution in middle latitudes and in the upper stratosphere. Several studies have addressed the significance of contributions of dynamical processes to long-term ozone changes (Appenzeller et al., 2000; Weiss et al., 2001; Wohltmann et al., 2007; Mäder et al., 2007; Rieder et al., 2010; Frossard et al., 2013; Rieder et al., 2013; Nair et al., 2013). Several papers (Harris et al., 2008; WMO, 2007) found that a significant part of the change in total ozone after the middle of the 1990s and until 2005 was caused by dynamical variability. Therefore, it is important to assess the contribution of dynamical parameters in order to derive ODS-associated

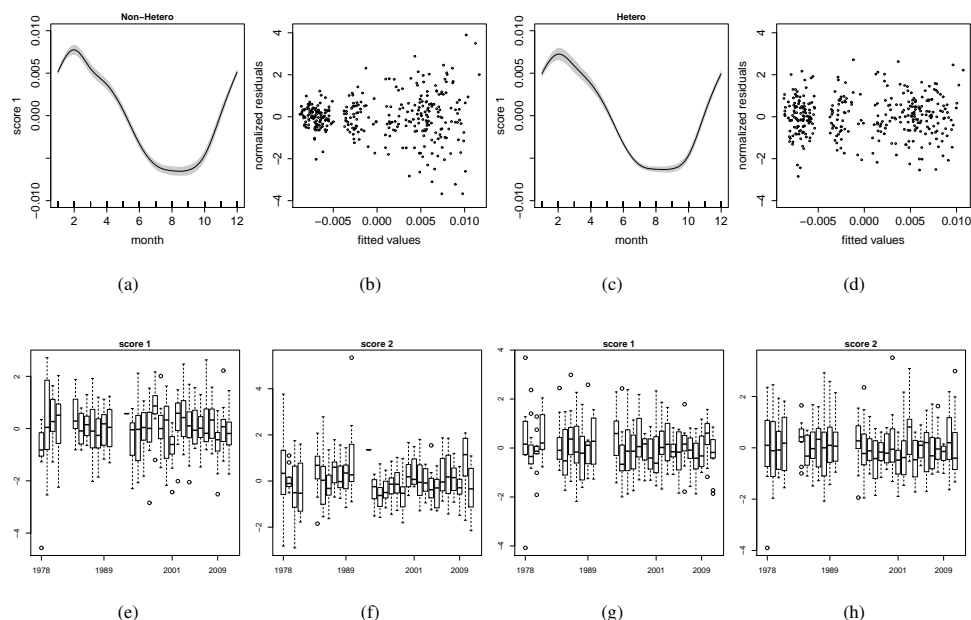


Fig. 10. (a) The fitted month curve (PC score 1 for Boulder) and its 95 % Bayesian confidence intervals when constant distribution is assumed for the error. (b) The fitted PC score 1 versus normalized residuals from the AMMs with the constant distribution assumption for Boulder. (c) The fitted month curve (PC score 1 for Boulder) and its 95 % Bayesian confidence intervals with the heteroscedastic error assumption. (d) The fitted PC score 1 versus normalized residuals from the AMMs with the heteroscedastic error assumption for Boulder. (e) Box plots of normalized residuals grouped by year: for Boulder, (e) and (f), and for Arosa, (g) and (h).

trends in Umkehr ozone profiles. Figures 11 and 13 show the time series plots of the monthly ozone profiles (Umkehr and estimated) and thus they have large seasonal variability. To show how well the time series are fitted by the model for each year, we display them as annual mean data (compute the annual mean of the Umkehr and estimated ozone profiles) in Figs. 12 and 14. They will help us show how the covariates contribute to ozone variability from one year to another (see Figs. 8 and 9).

We compute the monthly contribution of each covariate to changes in ozone profiles, and plot these monthly time series of % contributions as box plots by year in Figs. 8 and 9. The QBO2 (associated with winds at 10 hPa and 50 hPa) contributions at Boulder and at Arosa are very different, as seen in the second and third row of Fig. 8. For example, at 150 hPa this is a positive contribution up to 1 % at Boulder, while at Arosa there is no contribution from QBO2 at this pressure level. The opposite occurs at 16 and 8 hPa, where QBO2 does not affect ozone at Boulder, but affects ozone at Arosa (mostly a positive effect up to 1.5 %). In addition, in the Boulder data we see that QBO2 mostly decreases ozone at 53 and 16 hPa, while this is not the case at Arosa. QBO1 has a relatively large negative effect (up to -4%) on ozone at 253 hPa level in Boulder (see the first row of Fig. 8), but the effect is slightly smaller (up to -3%) at 150 hPa level; in the middle stratosphere (50–30 hPa) the response is mostly positive (up to 4 %).

ENSO does not show any significant contribution to Arosa ozone levels. However, it plays a significant role in modulating ozone below 53 hPa (up to -4%) for the Boulder data; see the bottom row of Fig. 8. This is expected, as the influence of ENSO should be larger over North America than over Europe. Also, it is interesting to see that since 2004 there is general negative trend in ozone forced largely by ENSO. Positive ENSO signals in 1987 (for all months) and 1997–1998 (for most months) are observed at Boulder as a positive contribution to ozone levels. Unfortunately, we do not see a positive ENSO effect on Umkehr ozone variability in 1983, since it was coincident with large aerosols remaining in stratosphere after El Chinchón eruption in 1982, which caused errors in Umkehr ozone retrieval. This is why we had to remove the Umkehr records during the 1982–1983 period. A negative phase of the ENSO index (La Niña) was observed in 1989, 1999/2000 and 2010/2011, and led to a large reduction in Boulder ozone. The apparent decline in Boulder ozone since 2004 is most likely related to the effects of ENSO on ozone variability. Arosa does not show any sensitivity to ENSO. This may be one of the reasons for the observed differences in middle stratospheric ozone trends over the 2000s at Boulder and Arosa.

AO time series feature a positive phase in 1989–1995, and again in 2007–2009, while a large negative excursion occurs in 2001 and 2010. Since 2000 the year-to-year variability in the AO has been found to be very large compared to

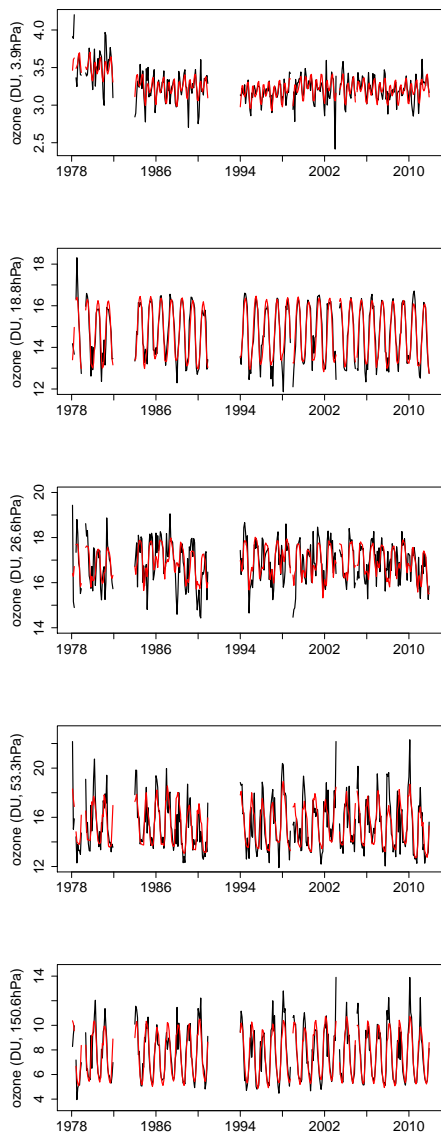


Fig. 11. Umkehr monthly ozone profiles (in black) and the estimated monthly profiles using the fitted PC scores from the AMM (in red) of Boulder. Volcanic years are removed. The ozone profiles are in DU.

the 1980s (mostly negative) and 1990s (mostly positive). In Boulder, the AO contributions are significant at 253 hPa (up to 3 % increased ozone), at 90 hPa and 50 hPa (decrease up to -2% and -1% respectively). Similarly, in Arosa, at altitude levels of 253–53 hPa, ozone is significantly influenced by AO (up to $\pm 4\%$). However, the AO contributions to ozone variability in the low and middle stratosphere are smaller at Boulder compared to Arosa, and this can be explained by the fact that AO brings stronger dynamical variability to higher latitudes.

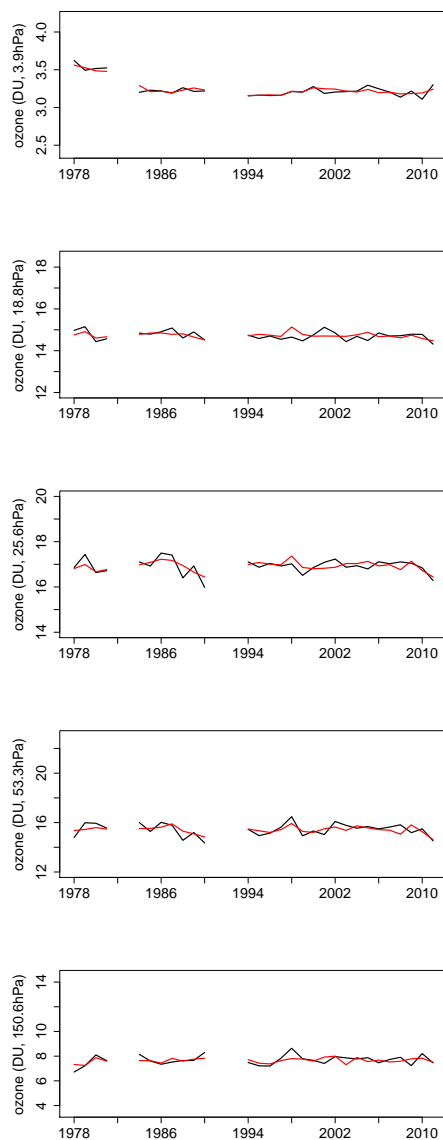


Fig. 12. Umkehr yearly ozone profiles (in black) as mean values of the Umkehr monthly ozone profiles in Fig. 11, and the estimated yearly profiles (in red) as mean values of the estimated monthly ozone profiles Fig. 11 of Boulder. Volcanic years are removed. The ozone profiles are in DU.

5.2 Residual analysis

Comparing the residual plots from AMMs with and without the heteroscedastic error structure is appropriate for model checking. Figure 10b gives normalized residuals versus fitted PC score 1 of Boulder when constant variance is assumed, while Fig. 10d demonstrates the same plot when heteroscedasticity is accounted for. In Fig. 10d, the residuals are evenly distributed around zero compared to Fig. 10b, where they are not. This supports the need for a more complex error structure in the model and the appropriateness of the

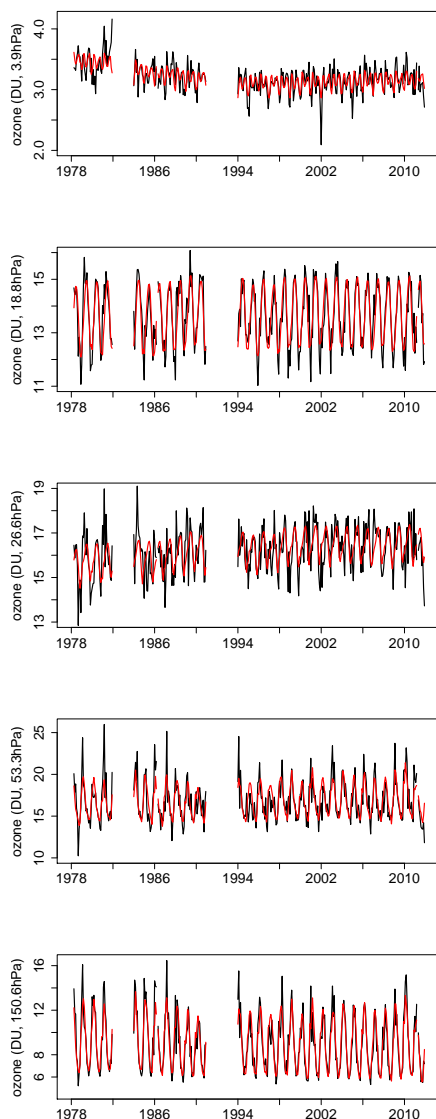


Fig. 13. Umkehr monthly ozone profiles (in black) and the estimated monthly profiles using the fitted PC scores from the AMM (in red) of Arosa. Volcanic years are removed. The ozone profiles are in DU.

variance function we proposed. Note that in Fig. 10b the normalized residuals are obtained by $\tilde{\epsilon}_{li} = (\theta_{li} - \hat{\theta}_{li}) / \hat{\sigma}_l$, whereas in Fig. 10d they are defined by

$$\tilde{\epsilon}_{li} = \frac{\theta_{li} - \hat{\theta}_{li}}{\hat{\sigma}_l \sqrt{\text{Var}(\epsilon_{li})}}. \quad (12)$$

The estimated variance of errors $\hat{\sigma}_l^2$ is summarized in Table 3.

As mentioned earlier, there is a clear improvement upon the assumption of constant variance of errors, as periodic patterns in residuals have disappeared. Normalized residuals computed by Eq. (12) were grouped by year, and we plotted them as box plots. We only display the box plot of PC

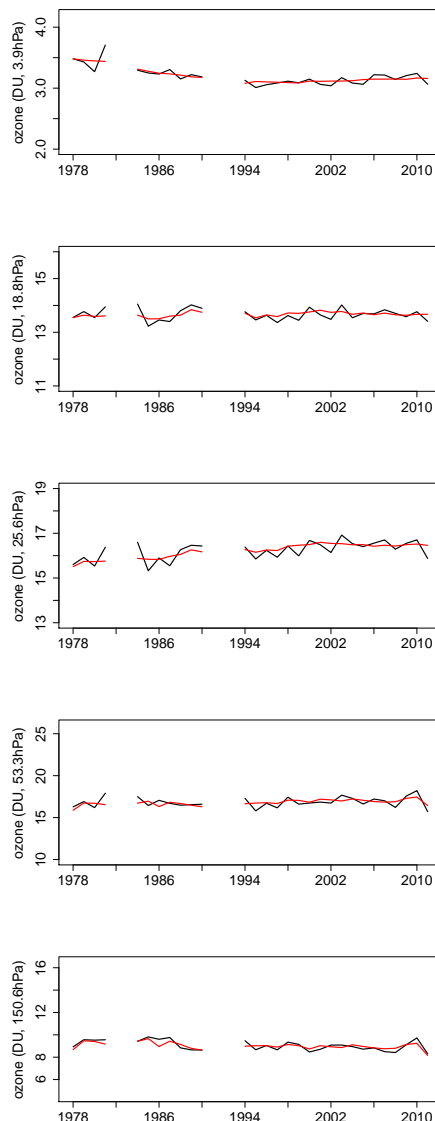


Fig. 14. Umkehr yearly ozone profiles (in black) as mean values of the Umkehr monthly ozone profiles in Fig. 13, and the estimated yearly profiles (in red) as mean values of the estimated monthly ozone profiles Fig. 13 of Arosa. Volcanic years are removed. The ozone profiles are in DU.

scores 1 and 2 in Fig. 10e–h. Large absolute values of residuals tend to occur in the early years when the measurements are considered to be less reliable or noisy. Indeed, the beginning of the Boulder record has fewer measurements and the observations were taken in manual mode (thus prone to operator errors). For Arosa, observational methods were established back in the 1960s and were not changed until the 1980s. Also, a smaller number of measurements per Umkehr curve was collected prior to the automation of the measurements system in 1986. The abrupt changes in the size of the normalized residuals in PC score 1 of Boulder data are found

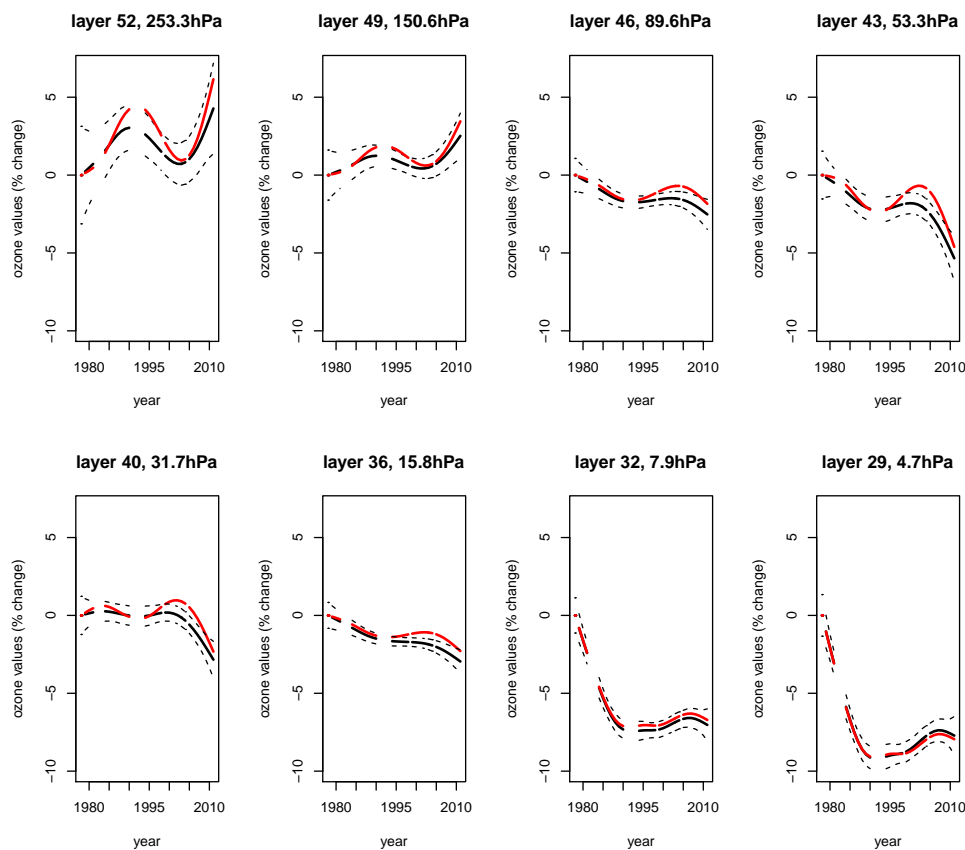


Fig. 15. Estimated ozone trends as % changes in Boulder at selected layers, with 95 % confidence intervals. The trends from the full regression model are in black solid lines together with their 95 % confidence intervals in dotted lines. The trends from the regression model without EP flux are in red solid lines. Volcanic years are omitted from the trend analysis.

in 2001; see Fig. 10e. PC score 1 captures atmospheric variability between 250 and 68 hPa. This feature could be related to the mechanism that caused an abrupt decrease in the stratospheric water vapor observed over the tropics in 2001 (Randel et al., 2006; Solomon et al., 2010). The changes were related to an increase in tropical upwelling and a change in the transport patterns. Note that a decrease in the low stratospheric water vapor was also observed at around the same time over Boulder (Hurst et al., 2011). Finally, in PC score 1 for Boulder, we might detect the influence of the ENSO cycle (supported by the fact that the EDF of ENSO is 1.7) – negative in 2001 (for all months) and positive in 1998 (for most months) in the residuals. Indeed, ENSO influences transport patterns and positively affects lower and middle stratospheric ozone at the northern middle latitudes.

6 Estimated ozone trends

Here we focus on the ozone trend analysis. In Sect. 6.1 we derive the trends for selected altitude levels using a functional approach. In Sect. 6.2 we also carry out an analysis of total

column ozone trends from two sources for the purpose of comparison. Finally, in Sect. 6.3 we investigate the effect of EP flux on ozone, as the EP flux is found to be the most significant proxy for explaining ozone variations.

6.1 Estimated trends for profiles using a functional approach

From AMMs we estimated the ozone trends for each PC score, and the estimation results are discussed in Sect. 5. Here, to derive ozone trends that take all PC scores into account after other effects of covariates (month, QBO, the solar cycle, AO, ENSO and EP flux) are removed, we compute the trend, for a given altitude level x , by

$$O_i(x) = \sum_{l=1}^5 \hat{\xi}_l(x) \hat{g}_{l2}(\text{Year}_i). \quad (13)$$

$O_i(x)$ is the estimated ozone at altitude x and year i , \hat{g}_{l2} are the fitted scores (only for the year term) from AMMs, and $\hat{\xi}_l(x)$ are the smoothed PCs as in Eq. (4) and are computed by Eq. (A1). Estimated ozone trend curves (in black solid lines)

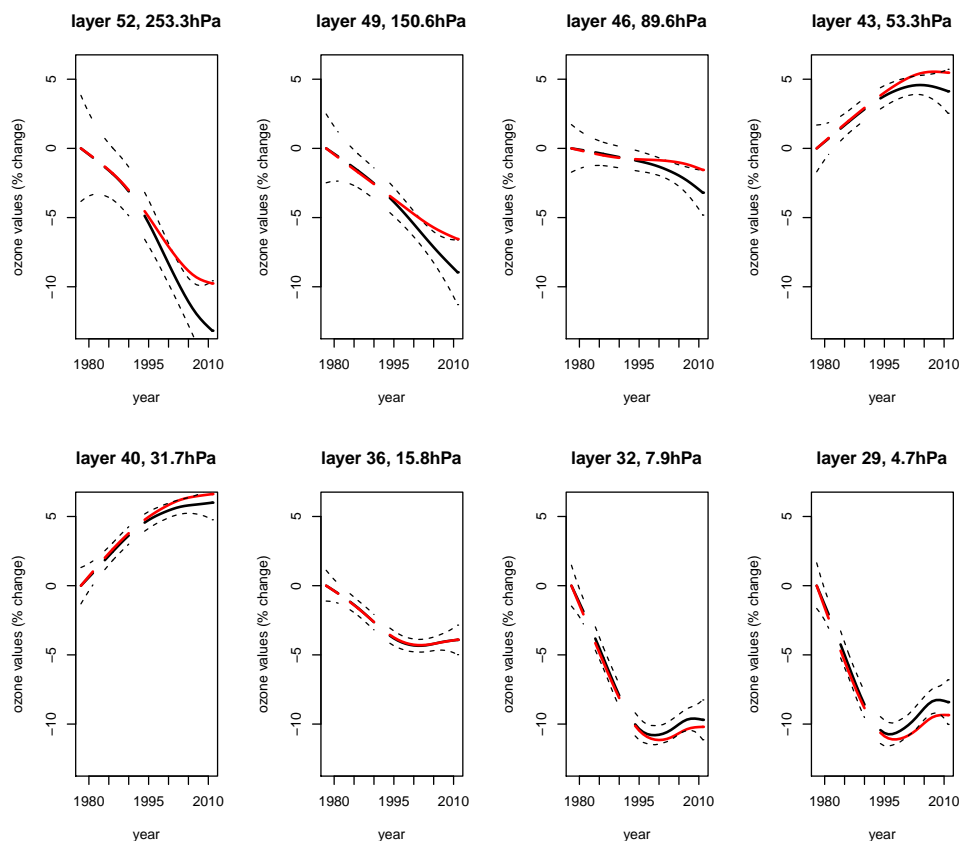


Fig. 16. Estimated ozone trends as % changes in Arosa at selected layers, with 95 % confidence intervals. The trends from the full regression model are in black solid lines together with their 95 % confidence intervals in dotted lines. The trends from the regression model without EP flux are in red solid lines. Volcanic years are omitted from the trend analysis.

with their 95 % confidence intervals (in dotted black lines) for Boulder (Fig. 15) and Arosa (Fig. 16) at selected altitudes are reported. For Boulder, the trends seem to show a typical decrease and a beginning of a recovery from 1996 onwards, but at 32–8 hPa from 2003, there is a sharp decline in estimated ozone trends (possibly attributed to the ENSO effect). Note that the fitted trend curve of the PC score 3 largely contributes to the sharp decline (the fitted trend curve associated with PC 3, see Fig. 6, displays a sharp decline after 2003). There is also a detected increase in the troposphere, as well as a leveling off in the upper stratosphere.

For Arosa, the very same PC decomposition as for Boulder does not yield similar trends. The different trend estimates might be related to various aspects. Nonlinear synoptic waves can affect stratospheric ozone through vertical transport on time scales that are shorter than photochemical lifetime, and can produce sufficient contributions to long-term variability on regional scales – meaning that there will be a difference between ozone variability at Boulder or Arosa (Hood and Soukharev, 2005). Hood et al. (1999) and Hood and Soukharev (2005) estimated that 40 % of the middle latitude ozone trend can be attributed to the increase in the

transient Rossby wave breaking frequency between 1979 and 1998 during the month of February. It is important to quantify dynamical transport associated with changes in atmospheric circulation that affects long-term ozone changes in comparison to chemically driven changes in ozone. For example, Hood and Soukharev (2005) also discussed that the pole-ward and equator-ward horizontal transport at middle latitudes (as predicted by the EP flux variability and is related to planetary wave forcing and changes in adiabatic Brewer–Dobson circulation) can be considered zonally averaged and should contribute similarly to Boulder and Arosa ozone variability. According to Hood and Soukharev (2005), the impact of the EP flux on ozone trends should increase with latitude, so it should be more significant at Arosa as compared to Boulder.

Let us discuss the possible impact of the solar radiation. Several papers recently published analyses of stratospheric ozone and temperature changes, related to the spectral output of the solar radiation during maximum and minimum of the last solar cycle 23 (Haigh et al., 2010; Oberländer et al., 2012). Solar radiation measurements made by the solar Stellar Irradiance Comparison Experiment (SOLSTICE)

and the Spectral Irradiance Monitor (SIM) instruments on the *SORCE* satellite indicate a spectral dependence in the UV and visible part of the solar spectrum observed between 2004 and 2007. The radiative chemical transport models were used to investigate effects of the solar spectrum perturbations on photolysis rates and the temperature of the upper and middle stratosphere. The change in ozone at 4 hPa level over northern middle latitudes between solar maximum and minimum was increased from 0.7 % to 1.8 % when the SIM data were used in the model simulation as compared to the climatological solar spectrum results. The attribution of the solar cycle in the regression model fitted to the *Aura* MLS (Microwave Limb Sensor) ozone time series was found to contribute about 4 % decrease in ozone at 10–6.8 hPa pressure levels from 2004 to 2007. This effect is related to the spectral variability of cycle 23 and was not found in the previous cycles. The nonlinear fit of the PC score 4 of Boulder for solar cycle signal, see Fig. 6, indicates that the AMM model could not produce the best fit for the solar cycle in ozone data at the middle stratosphere and troposphere (score 4). This could be related to the changes in the spectral distribution of the solar radiation over the last 3 solar cycles that might have created a different response in middle latitude ozone.

6.2 Estimated trends for total column ozone

We included an analysis of total column ozone from two sources (WOUDC published total ozone, and Umkehr total ozone by summing up all layers across the profile). It only uses the second step (i.e., the AMMs on the covariates: the only possible common solution to the trend analysis problem for profile and total column data sets) of our statistical analysis to derive trends as well as influences of covariates. Total column ozone is the integral of the profile, and therefore the influences of covariates are somehow integrated along the profile when carrying out the regression: this can result in a loss of accuracy whenever a covariate's influence varies with altitude, as we noticed in our full analysis for, e.g., solar or EP flux. So, we emphasize that the total column ozone analysis may not yield as precise outcomes as a full profile approach.

For Boulder, the trends derived from the two total column data sets differ in the middle of the analyzed period (1990–2005); see Fig. 17a. It creates almost a linear trend in Umkehr total ozone (TO) data (trend in red solid line and the 95 % confidence intervals in red dotted line), while the WOUDC data set suggests very strong decline in TO until 1996 and then a very slow recovery until 2011 (trend in black solid line and the 95 % confidence intervals in black dotted line). When the ENSO signal is not used as an explanatory parameter, the WOUDC trend becomes more negative at the end of the time series and more in agreement with the Umkehr TO-based trend; see Fig. 17b.

For Arosa, we find that for total column ozone, the trend is negative over the full period 1978–2011; see Fig. 18a. When

carrying out our profile analysis, the trend is also negative for altitudes below approximately 63 hPa, but can be positive for altitudes slightly above 63 hPa over 1978–2011; see the black solid lines in Fig. 16. Rieder et al. (2010) also found a negative TO trend at Arosa (see their Fig. 4 for 1978–2008). From our analysis at Arosa, using WOUDC data, we first notice that our trends are approximately linear; see Fig. 18b. The linear trend (and associated standard error) for 1980–1990 is -1.2% ($\pm 0.4\%$), for 1990–2000 it is -1.3% ($\pm 0.4\%$), and for 2000–2010 it is -1.1% ($\pm 0.4\%$). The linear trends for TO at Arosa from Rieder et al. (2010), see their Table 3, are not as negative as those found here. For the only overlapping period of 1980–1990, the trend is -0.9% ($\pm 0.3\%$) per decade when extreme events are excluded. Note that the trends from Rieder et al. (2010) are much higher when all observations are kept, at -2.4% ($\pm 0.5\%$) per decade. We can now discuss the differences between our analysis and theirs. Our method does not take out extreme events. In Rieder et al. (2010), many of the extreme events are associated with volcanic eruptions, whereas we remove two entire volcanic periods from the start. Other extreme events are associated with ENSO, NAO, the amount of ODSs and the strength of the Polar vortex. Although this study did not account for ODSs, the influence of ENSO (not significant in Arosa, while significant in the Boulder analysis) and AO (significant in both data sets) is attributed in a regression setting, in contrast to the approach of Rieder et al. (2010) that excluded extreme events. We also accounted for solar and QBO effects, unlike Rieder et al. (2010), but used EP flux, not Polar vortex ozone loss contribution to northern mid-latitudes, as the dynamical contribution.

Rieder et al. (2010) also studied the occurrence of the extreme high and low events in the Arosa TO record. Results of their analysis indicate that since the 1990s, the Polar vortex has strengthened, which also creates a lower TO column at high latitudes (Chipperfield and Jones, 1999; Hadjinicolaou et al., 2002). The evidence for strengthening of the vortex is supported by the increase in the negative vortex index from the 1980s to 2000; see Fig. 1 in Rieder et al. (2010). Results of Rieder et al. (2010) imply that depleted ozone after the break of the Arctic Polar vortex is mixed into middle latitudes, which dilutes high spring ozone, thus the frequency of extreme high ozone events at Arosa is reduced in the last two decades. This effect likely produces more negative trends in TO at Arosa. Since these extreme values are strongly associated with dynamics (Rieder et al., 2010; Frossard et al., 2013; Rieder et al., 2013), and we aim to account for these effects through the use of AMMs on the covariates, our trend analysis of TO over Arosa is in good agreement with Rieder et al. (2010), though not fully due to the differences in the approaches.

Rieder et al. (2010) mention the sensitivity of total column ozone variability over Arosa to NAO and ENSO events, such as a decrease in TO associated with the positive phase of NAO. Rieder et al. (2010) also mention many more positive

NAO events after 1990 (1992–1993, 1995, 2000, 2005, 2007 and 2008) as compared to negative NAO events (only 1996). It means that the frequency of these events has changed with time. This may result in an increase in extreme low TO values in the recent years (after 1990). Indeed, the positive NAO events are associated with decreased ozone over Europe. Therefore, after 1990, predominantly the positive NAO events would create a tendency for ozone trends to become more negative in Arosa, in contrast to the tendency for TO to increase due to reduction in ODSs.

The EP flux proxy plays an important role in deriving trends for both the WOUDC and Umkehr records (of Arosa) by making the 1997–2011 trend more negative; see Fig. 17c and d and Fig. 18b and c. Without using the EP flux proxy, the derived trend in Boulder Umkehr TO data (see Fig. 17(d)) shows an increase in ozone from 1995 to 2006, with a subsequent decrease through 2011. Note that our profile analysis results in trends that are of various shapes for Boulder (see the black solid lines in Fig. 16), so they cannot readily be compared with TO trend analyses.

6.3 The effect of EP flux on ozone trend

The EP flux has the greatest influence for most scores, so it is the most influential covariate for explaining ozone variability. To see its effect on the estimated trends, we compare trends from the full regression model (including covariates of month, year, QBO1, QBO2, solar, AO, ENSO and EP flux) in black solid lines with the trends from the regression model without EP flux in red solid lines; see Figs. 15 (Boulder) and 16 (Arosa).

We find that adding EP flux to the explanatory parameters changed the long-term trends in Boulder Umkehr data in two different ways. First of all, we found a difference in the trends in the middle of time series (1985 and 2000) when analyzing trends with and without EP flux at altitude between 360 and 105 hPa. At the same time, at altitudes between 15 and 45 hPa, we found an increased difference in trends at the end of time series (2000 to 2011). However, we found no significant effect of the EP flux contribution on ozone trends at higher altitudes (15–2 hPa), where transport plays a less important role as compared to the lower and middle stratosphere.

We find that use of the EP flux in trends analysis of the Arosa Umkehr data at altitudes above 32 hPa produces similar effects to what we found for Boulder. It appears that in the upper stratosphere the effect of EP flux on Arosa trends is slightly greater than for Boulder. In the upper stratosphere, trends tend to be less negative in the last 10 yr if EP flux is used in the regression. Thus, we can say that upper stratospheric ozone recovery (above 8 hPa or 35 km) in Arosa is more pronounced when we attribute some of the recent 10 yr of ozone changes to the EP flux variability. However, in the middle stratosphere (16–64 hPa) in Arosa, we do not find negative trends as compared to trends for Boulder. The con-

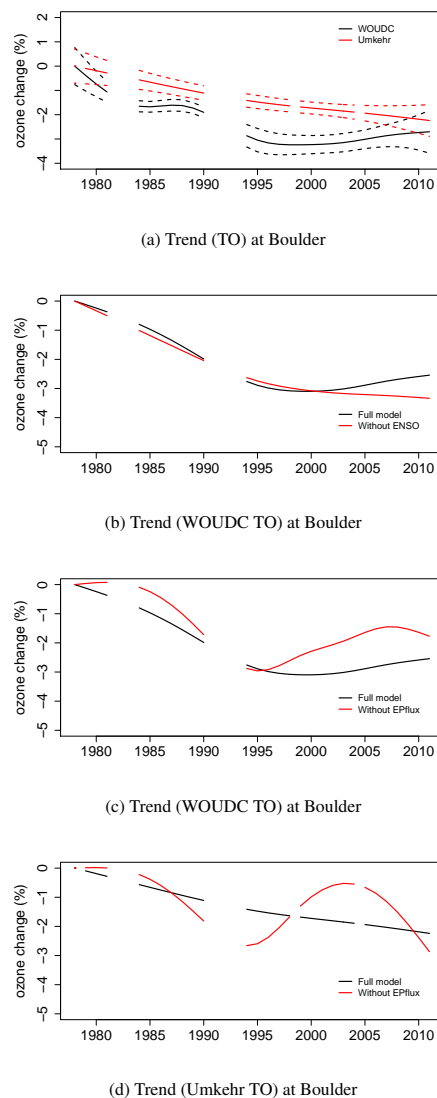


Fig. 17. Estimated trends for total column ozone at Boulder. **(a)** Ozone trends for Umkehr (in red solid line) and WOUDC (in black solid line) TO, with their 95% confidence intervals in dotted lines with the matching color. **(b)** Ozone trends for WOUDC TO obtained from the full regression model (on black) and the model without ENSO (in red). **(c)** Ozone trends for WOUDC TO obtained from the full regression model (in black) and the model without EP flux (in red). **(d)** Ozone trends for Umkehr TO obtained from the full regression model (in black) and the model without EP flux (in red).

tribution to the trend estimation at these altitudes of the EP flux is significant and makes the trends less positive (however, it does not make trends negative). Note that from the plots for WOUDC TO at Arosa, see Fig. 18b, the removal of the EP flux covariate reduced the trend from -4 to -3 % when EP flux was removed. At lower altitudes, trends in Arosa start to change and become negative below 64 hPa (below 20 km), whereas in Boulder trends become positive in the lower stratosphere. The Arosa analysis shows a large

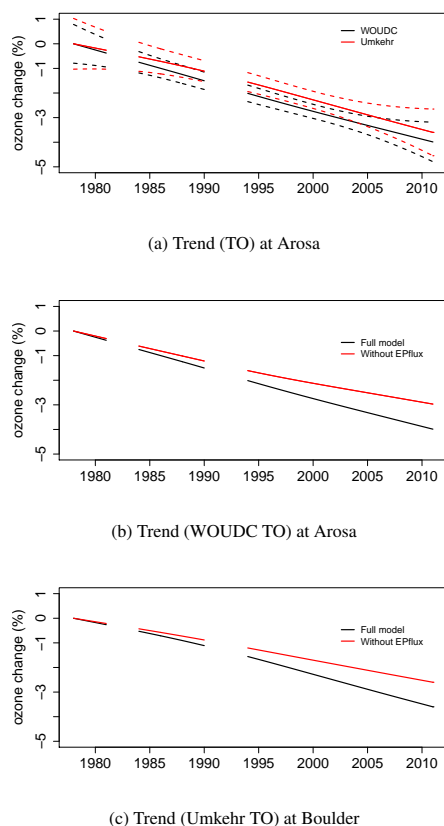


Fig. 18. Estimated trends for total column ozone at Arosa. **(a)** Ozone trends for Umkehr (in red solid line) and WOUDC (in black solid line) TO, with their 95 % confidence intervals in dotted lines with the matching color. **(b)** Ozone trends for WOUDC TO obtained from the full regression model (in black) and the model without EP flux (in red). **(c)** Ozone trends for Umkehr TO obtained from the full regression model (in black) and the model without EP flux (in red).

contribution of the EP flux in the lower stratosphere and troposphere (below 20 km) and makes resulting trends even more negative. This is consistent with the fact that the TO derived from Arosa WOUDC and Umkehr data continues to decrease after 1996. It is not clear what other processes might have contributed to these changes over the last 10 yr. We only account for the large-scale transport contribution toward the Arosa trends. There might be additional synoptic scale fluctuations that create recent negative changes in ozone in the troposphere and the lower stratosphere, and therefore these are observed in the Arosa TO trends (Steinbrecht et al., 2011).

7 Conclusions

The recovery of ozone is strongly related to the remaining concentration of equivalent effective stratospheric chlorine (EESC). Recently, several publications revised lifetimes of

the CFCs and other ODSs in the stratosphere (Laube et al., 2012). The age of air terminology is used to predict how much of the EESC remains in the stratosphere after the ODS are released at the surface. The quality of prediction is related to the uncertainties in estimates of the transport of chemicals and the fractional release factors of the chlorine in the stratosphere. The age of air related to the time for transport from troposphere to stratosphere and then to the higher latitudes from tropics has been reassessed based on the changes in the Brewer–Dobson circulation. Recently, more emphases have been put on increases in the HydroChloroFluoroCarbon (HCFC) concentrations that are used to replace CFCs under the Montreal Protocol. The HCFCs are one class of chemicals being used to replace the CFCs. They contain chlorine and so deplete stratospheric ozone, but to a much lesser extent than CFCs. However, HCFCs have a large global warming potential and thus cannot be ignored.

The use of AMMs, with seasonal variance, for ozone profiles enables better quantification of uncertainties than previous methods. It thus gives us more confidence in the estimates of the influences of the various variables. In particular, the trends over the years displayed for the few modes of variability in the atmosphere identified by the functional PCs show interesting and significant behaviors. As a result we are able to identify features present in the data that could be associated with the semi-annual oscillation; this reinforces our faith in the statistical model. Furthermore, the statistical detection of a recovery is more likely to occur using methods that treat influences from various factors in a nonlinear fashion, adjust for heteroscedasticity, and consider modes of variability across altitudes for more robust results. Nevertheless, one may consider adding more covariates such as stratospheric water vapor to pin down the trends more precisely. That increase in complexity of the model might not however result in a significant improvement. One may also consider removing extreme values – to be defined possibly as large values of the scores – as these can have a large influence on trends (Rieder et al., 2010) but be associated with dynamical features, not a genuine long-term evolution. However, great care in the selection of these values is necessary: see the discussion about exceedances over a threshold and r -largest order statistics models in Frossard et al. (2013).

For further study, we could specify a model that includes a latitude (and possibly longitude) argument as a covariate. As a result, we could borrow strength across the stations to improve the estimation of the influence of covariates. Indeed, the spatial fingerprints of the dynamical and chemical contributions to ozone variations (Frossard et al., 2013; Rieder et al., 2013) have highlighted such coherence (albeit only for total column ozone). An advanced formulation in this direction may require building spatial covariance models on the sphere (see, e.g., Bolin and Lindgren (2011) or Jun (2011) for multivariate spatial processes if one wants to consider other chemical species jointly). As a result, one could

potentially obtain a clearer picture of the evolution of ozone levels throughout altitudes on a global scale.

Another improvement would be to include short-term dynamical terms to potentially better fit scores 3 and 4. And one possibility would be to focus on extreme values of these scores, since for total column ozone, the relationship between long-term dynamical terms is stronger for extreme values than for mean values (Frossard et al., 2013), and is likely to be even stronger for short-term dynamical terms. Another possibility would be to use reanalysis data to see if dynamical transport information could yield a larger explanatory power for the AMM. However, such an effort would require some adequate dimension reduction of the dynamical terms to be able to include these. One approach might be to let a numerical model compute the effect of such short-term dynamical terms only (Guillas et al., 2004; Kobayashi and Shibata, 2011).

Appendix A

Statistical details

A1 FPCA

Silverman (1996) proposed a method of estimating smooth PCs by incorporating a penalty term into the orthonormality constraint imposed on the PCs. Further, following Ramsay and Silverman (2005, Ch. 8–9), we used a B-spline basis to represent the PCs, e.g., $\xi_l(x) = \sum_{k=1}^K z_{lk} \phi_k(x)$ and measured its roughness by the integrated squared second derivatives of the function. Then the roughness of each PC is easily quantified by using approximation to the integral, i.e., $\sum_{j=1}^{32} \sum_{k=1}^K [z_{lk} D^2 \phi_k(x_j)]^2$. We placed the knots at each layer, excluding the two end layers 29 and 60, and used cubic B-splines, thus $K = 32$. After projecting the PCs onto a B-spline basis, the fitting criterion for the coefficient vectors \mathbf{z}_l of the PCs becomes

$$\max \frac{\mathbf{z}_l^T \Phi^T \hat{\mathbf{Y}}^T \hat{\mathbf{Y}} \Phi \mathbf{z}_l}{\mathbf{z}_l^T (\mathbf{I}_K + \lambda \mathcal{P}^T \mathcal{P}) \mathbf{z}_l}, \quad (\text{A1})$$

where we denote $\hat{\mathbf{Y}} \in \mathbb{R}^{n \times K}$ the matrix consisting of (i, j) entry $\hat{y}_{ij} - \frac{1}{n} \sum_{i=1}^n \hat{y}_{ij}$, the centered fitted values of ozone from the smoothing splines, see Eq. (2), and denote $\Phi \in \mathbb{R}^{32 \times K}$ the B-splines basis matrix composed of (j, k) entry $\phi_k(x_j)$. $\mathcal{P}^T \mathcal{P}$ is the penalty matrix and $\mathbf{z}_l^T \mathcal{P}^T \mathcal{P} \mathbf{z}_l$ quantifies the roughness of each PC. The smoothing parameter λ controls the level of smoothing. We applied the same smoothing parameter to obtain all PCs. We selected $\lambda = 10^{-5}$, the same number used in fitting the smoothing spline models for Umkehr ozone profiles. Note that the numerator is the objective function to be maximized and that the denominator is the orthonormal constraint imposed on the PCs. Once $\hat{\mathbf{z}}_l$ are estimated from Eq. (A1), we obtain the PC by

$\hat{\xi}_l(x) = \sum_{k=1}^K \hat{z}_{lk} \phi_k(x)$. The R-function `pcafd` in the FDA library is used to compute the coefficients z_{lk} .

A2 Representation of penalized regression splines as additive terms

Here we review additive models where each term is represented by the penalized regression cubic spline. For each smooth term g_{lj} related to covariate F_j , for $j = 1, \dots, 8$, we have the model matrix of F_j denoted by $\mathbf{M}_j \in \mathbb{R}^{n \times K_j}$, where n is the number of samples and K_j is the number of basis functions for covariate F_j without a constant basis. Denoting \mathbf{M}_{ij} the i th row vector of the model matrix \mathbf{M}_j , we have $\mathbf{M}_{ij} = [F_{ij}, F_{ij}^2, F_{ij}^3, (F_{ij} - s_{1+})_+^3, \dots, (F_{ij} - s_{H+})_+^3]$, where $(F - s_h)_+ = \max\{0, (F - s_h)\}$ with knots s_h (Ruppert et al., 2003). Stacking up all smooth terms and an intercept, we define the i th row of the model matrix $\mathbf{B} \in \mathbb{R}^{n \times (K_\alpha + 1)}$, with $K_\alpha = \sum_{j=1}^8 K_j$, by $\mathbf{B}_i = [1, \mathbf{M}_{i1}, \dots, \mathbf{M}_{i8}]$. Denoting the coefficient vector of each smooth term by $\boldsymbol{\alpha}_{lj} = [\alpha_{lj1}, \dots, \alpha_{ljK_j}]$ and stacking up all coefficient vectors including an intercept, i.e., $\boldsymbol{\alpha}_l = [c_l, \boldsymbol{\alpha}_{l1}, \dots, \boldsymbol{\alpha}_{l8}]^T$, we minimize the following penalized least squares fitting criterion with respect to $\boldsymbol{\alpha}_l$:

$$\min \left\{ \|\boldsymbol{\theta}_l - \mathbf{B} \boldsymbol{\alpha}_l\|^2 + \boldsymbol{\alpha}_l^T \mathbf{S}_l \boldsymbol{\alpha}_l \right\}. \quad (\text{A2})$$

Here $\boldsymbol{\theta}_l = [\theta_{l1}, \dots, \theta_{ln}]^T \in \mathbb{R}^n$ is the l th score vector. The penalty matrix \mathbf{S}_l is a block diagonal matrix with blocks $\lambda_{lj} \mathbf{A}_j$:

$$\mathbf{S}_l = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \lambda_{l1} \mathbf{A}_1 & & \\ 0 & & \lambda_{l2} \mathbf{A}_2 & \\ \vdots & & \vdots & \vdots \\ 0 & & & \lambda_{l8} \mathbf{A}_8 \end{bmatrix} \quad (\text{A3})$$

where \mathbf{A}_j is the penalty matrix corresponding to the covariate F_j , which consists of the squared second derivatives of the basis functions we use. Note that the second derivatives of linear terms are zero, thus the first row and column of \mathbf{A}_j are all zero (penalty null space), which means that the linear terms are not penalized. λ_{lj} is the smoothing parameter that controls the smoothness of the function g_{lj} . Given λ_{lj} , the penalized least squares estimator of $\boldsymbol{\alpha}_l$ can be written as

$$\hat{\boldsymbol{\alpha}}_l = (\mathbf{B}^T \mathbf{B} + \mathbf{S}_l)^{-1} \mathbf{B}^T \boldsymbol{\theta}_l. \quad (\text{A4})$$

A3 Penalized regression splines as mixed effects models

We compare the penalized regression splines with the mixed effects models to understand how the variance–covariance matrix $\boldsymbol{\Omega}_{\lambda_l}$ achieves smoothing. In other words, we show that in AMMs the idea of penalization is incorporated into the variance–covariance matrix of a random vector via a Bayesian approach.

Denoting $p(X)$ the density function of a random variable X , we impose some prior beliefs on the coefficient vector of

the random effects, i.e., $\mathbf{u}_l \sim N(0, \mathbf{\Omega}_{\lambda_l})$; see Eq. (9); thus we have $p(\mathbf{u}_l) \propto \exp(-\frac{1}{2}\mathbf{u}_l^T \mathbf{\Omega}_{\lambda_l}^{-1} \mathbf{u}_l)$. Stacking up the coefficient vector of the fixed and random effects $\boldsymbol{\alpha}_l = [\mathbf{b}_l^T, \mathbf{u}_l^T]^T$, we define the inverse covariance matrix of the coefficient vector $\boldsymbol{\alpha}_l$ by

$$\mathbf{\Omega}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Omega}_{\lambda_l}^{-1} \end{bmatrix}, \quad (\text{A5})$$

where rows and columns of zero in the inverse matrix correspond to the variances or covariances of the coefficients related to the fixed effects. Then we write the density of the coefficient vector, i.e., $p(\boldsymbol{\alpha}_l) \propto \exp(-\frac{1}{2}\boldsymbol{\alpha}_l^T \mathbf{\Omega}^{-1} \boldsymbol{\alpha}_l)$, and have the conditional density of $\boldsymbol{\theta}_l$ given $\boldsymbol{\alpha}_l$:

$$p(\boldsymbol{\theta}_l | \boldsymbol{\alpha}_l) \propto \exp[-\frac{1}{2\sigma_l^2} (\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l)^T (\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l)]. \quad (\text{A6})$$

Using Bayes' rule, i.e., $p(\boldsymbol{\alpha}_l | \boldsymbol{\theta}_l) \propto p(\boldsymbol{\theta}_l | \boldsymbol{\alpha}_l) p(\boldsymbol{\alpha}_l)$, we compute the posterior density:

$$\begin{aligned} p(\boldsymbol{\alpha}_l | \boldsymbol{\theta}_l) &\propto \exp[-\frac{1}{2\sigma_l^2} (\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l)^T (\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l)] \\ &\quad -\frac{1}{2}\boldsymbol{\alpha}_l^T \mathbf{\Omega}^{-1} \boldsymbol{\alpha}_l] \\ &\propto \exp[-\frac{1}{2\sigma_l^2} ((\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l)^T (\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l) \\ &\quad + \sigma_l^2 \boldsymbol{\alpha}_l^T \mathbf{\Omega}^{-1} \boldsymbol{\alpha}_l)] \end{aligned} \quad (\text{A7})$$

Differentiating the posterior density with respect to the vector $\boldsymbol{\alpha}_l$, we have the equation

$$\mathbf{B}^T \boldsymbol{\theta}_l = (\mathbf{B}^T \mathbf{B} + \sigma_l^2 \mathbf{\Omega}^{-1}) \boldsymbol{\alpha}_l. \quad (\text{A9})$$

Therefore the Maximum Likelihood estimate of $\boldsymbol{\alpha}_l$ is given by

$$\hat{\boldsymbol{\alpha}}_l = (\mathbf{B}^T \mathbf{B} + \sigma_l^2 \mathbf{\Omega}^{-1})^{-1} \mathbf{B}^T \boldsymbol{\theta}_l. \quad (\text{A10})$$

We can compare the ML estimate in Eq. (A10) with the penalized least squares estimate in Eq. (A4). Then, we deduce $\sigma_l^2 \mathbf{\Omega}^{-1} = \mathbf{S}_l$. This implies that in the mixed effects model framework of the penalized regression splines, smoothing is induced by the two variance components, i.e., σ_l^2 and $\mathbf{\Omega}^{-1}$. Because in the penalty matrix \mathbf{S}_l we have known quantities of \mathbf{A}_j and the unknown parameters λ_{lj} , the estimation of the covariance matrix of the random effects $\mathbf{\Omega}_{\lambda_l}$ suffices to estimate the smoothing parameters λ_{lj} . Estimation of the variance components in the mixed effects model framework will be given in Appendix A4.

A4 AMM

In practice, $\mathbf{\Omega}_{\lambda_l}$ and σ_l^2 are unknown parameters, so we must estimate them. The estimation of $\mathbf{\Omega}_{\lambda_l}$ and σ_l^2 will follow.

Denoting $\text{Cov}(\boldsymbol{\theta}_l) := \boldsymbol{\Sigma}_l$, we deduce $\text{Cov}(\boldsymbol{\theta}_l) = \mathbf{B}_R \mathbf{\Omega}_{\lambda_l} \mathbf{B}_R^T + \sigma_l^2 \mathbf{I}_n$, under the Gaussian and the independence of $\boldsymbol{\epsilon}_l$ and \mathbf{u}_l assumptions. Then the BLUP (Robinson, 1991) of \mathbf{u}_l is given by

$$\hat{\mathbf{u}}_l = \hat{\mathbf{\Omega}}_{\lambda_l} \mathbf{B}_R^T \hat{\boldsymbol{\Sigma}}_l^{-1} (\boldsymbol{\theta}_l - \mathbf{B}_F \hat{\mathbf{b}}_l) \quad (\text{A11})$$

where $\hat{\mathbf{\Omega}}_{\lambda_l}$, $\hat{\boldsymbol{\Sigma}}_l$ and $\hat{\mathbf{b}}_l$ are the estimates of $\mathbf{\Omega}_{\lambda_l}$, $\boldsymbol{\Sigma}_l$ and \mathbf{b}_l . By incorporating the random components in Eq. (9) into the error term, i.e., $\mathbf{e} = \mathbf{B}_R \mathbf{u}_l + \boldsymbol{\epsilon}_l$, we write the model as

$$\begin{aligned} \boldsymbol{\theta}_l &= \mathbf{B}_F \mathbf{b}_l + \mathbf{e} \\ \mathbf{e} &\sim N(0, \boldsymbol{\Sigma}_l). \end{aligned} \quad (\text{A12})$$

Then the ML estimate of \mathbf{b}_l is computed by $\hat{\mathbf{b}}_l = (\mathbf{B}_F^T \hat{\boldsymbol{\Sigma}}_l^{-1} \mathbf{B}_F \mathbf{B}_F^T \hat{\boldsymbol{\Sigma}}_l^{-1} \boldsymbol{\theta}_l)$.

Estimation of \mathbf{b}_l and prediction of \mathbf{u}_l requires the estimate $\hat{\boldsymbol{\Sigma}}_l$, and the commonly used algorithms for estimation of the variance components are ML and REstricted Maximum Likelihood (REML). It is known that ML estimates of variance parameters tend to become seriously biased, as the number of fixed parameters in the model increases, because the ML method does not account for the degrees of freedom lost resulting from the fixed effects estimation (Wood, 2006). In the REML method, instead of maximizing the log-likelihood function with respect to the variance parameter, the average of the likelihood over all possible values of \mathbf{b}_l is maximized. More specifically, the log-likelihood function of $\boldsymbol{\theta}_l$

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}_l | \mathbf{b}_l, \boldsymbol{\Sigma}_l) &\propto -\frac{1}{2} [\log |\boldsymbol{\Sigma}_l| \\ &\quad + (\boldsymbol{\theta}_l - \mathbf{B}_F \mathbf{b}_l)^T \boldsymbol{\Sigma}_l^{-1} (\boldsymbol{\theta}_l - \mathbf{B}_F \mathbf{b}_l)] \end{aligned} \quad (\text{A13})$$

is replaced by the integral $\mathcal{L}_R(\boldsymbol{\theta}_l | \boldsymbol{\Sigma}_l) = \int \mathcal{L}(\boldsymbol{\theta}_l | \boldsymbol{\Sigma}_l) d\mathbf{b}_l$. For a more detailed discussion of the REML criterion, see, e.g., Harville (1977). The closed form of the integral can be written as

$$\mathcal{L}_R(\boldsymbol{\theta}_l | \boldsymbol{\Sigma}_l) = \mathcal{L} - \frac{1}{2} \log |\mathbf{B}_F^T \boldsymbol{\Sigma}_l^{-1} \mathbf{B}_F|, \quad (\text{A14})$$

where \mathcal{L} is in Eq. (A13). The restricted log-likelihood \mathcal{L}_R is not analytically maximized. Thus, numerical approximation (e.g., Newton–Raphson methods) is widely used. The `mgcv` package in R fits $\boldsymbol{\Sigma}_l$ by a moderate number of expectation–maximization (EM) iterations to refine the starting values followed by a Newton–Raphson optimization. The consequence of the REML is a more robust estimate of the misspecification of the error structure (Opsomer et al., 2001), and it avoids overfitting (Pinheiro and Bates, 2000).

Acknowledgements. Ah Yeon Park was supported by a doctoral fellowship from the Kwanjeong Educational Foundation. S. Guillas was partially supported by a Leverhulme Trust research fellowship on “stratospheric ozone and climate change” (RF/9/RFG/2010/0427). We also thank the NOAA Atmospheric

Composition and Climate program as the source of funding for this paper.

Edited by: S. Godin-Beekmann

References

- Appenzeller, C., Weiss, A. K., and Staehelin, J.: North Atlantic oscillation modulates total ozone winter trends, *Geophys. Res. Lett.*, 27, 1131–1134, 2000.
- Bolin, D. and Lindgren, F.: Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping, *Ann. Appl. Stat.*, 5, 523–550, 2011.
- Bosq, D.: *Linear Processes in Functional Spaces: theory and applications*, Springer, New York, 2000.
- Chipperfield, M. P. and Jones, R. L.: Relative influences of atmospheric chemistry and transport on Arctic ozone trends, *Nature*, 400, 551–554, 1999.
- Craven, P. and Wahba, G.: Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation, *Numerische Mathematik*, 31, 377–403, 1979.
- Crutzen, P. J.: Estimates of Possible Variations in Total Ozone Due to Natural Causes and Human Activities, *Ambio*, 3, 201–210, 1974.
- de Boor, C.: *A Practical Guide to Splines*, Springer, New York, 2001.
- Dhomse, S., Weber, M., Wohltmann, I., Rex, M., and Burrows, J. P.: 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, doi:10.5194/acp-6-1165-2006, 2006.
- Farman, J., Gardiner, B. G., and Shanklin, J. D.: Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction, *Nature*, 315, 207–210, 1985.
- Frossard, L., Rieder, H. E., Ribatet, M., Staehelin, J., Maeder, J. A., Di Rocco, S., Davison, A. C., and Peter, T.: On the relationship between total ozone and atmospheric dynamics and chemistry at mid-latitudes – Part 1: Statistical models and spatial fingerprints of atmospheric dynamics and chemistry, *Atmos. Chem. Phys.*, 13, 147–164, doi:10.5194/acp-13-147-2013, 2013.
- Garcia, R., Dunkerton, T., Lieberman, R., and Vincent, R.: Climatology of the semiannual oscillation of the tropical middle atmosphere, *J. Geophys. Res.*, 102, 26019–26032, 1997.
- Gillett, N. P., Akiyoshi, H., Bekki, S., Braesicke, P., Eyring, V., Garcia, R., Karpechko, A. Yu., McLinden, C. A., Morgenstern, O., Plummer, D. A., Pyle, J. A., Rozanov, E., Scinocca, J., and Shibata, K.: Attribution of observed changes in stratospheric ozone and temperature, *Atmos. Chem. Phys.*, 11, 599–609, doi:10.5194/acp-11-599-2011, 2011.
- Gu, C.: *Smoothing Spline ANOVA Models*, Springer, New York, 2002.
- Guillas, S., Stein, M., Wuebbles, D., and Xia, J.: Using chemistry transport modeling in statistical analysis of stratospheric ozone trends from observations, *J. Geophys. Res.-Atmos.*, 109, D22303, doi:10.1029/2004JD005049, 2004.
- Hadjinicolaou, P., Jrrar, A., Pyle, J. A., and Bishop, L.: The dynamically driven long-term trend in stratospheric ozone over northern middle latitudes, *Q. J. Roy. Meteorol. Soc.*, 128, 1393–1412, 2002.
- Haigh, J. D., Winning, A. R., Toumi, R., and Harder, J. W.: An influence of solar spectral variations on radiative forcing of climate, *Nature*, 467, 696–699, 2010.
- 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., Karpechko, A., Kivi, R., Knudsen, B., Krizan, P., Lastovicka, J., Maeder, J., Orsolini, Y., Pyle, J. A., Rex, M., Vanicek, K., Weber, M., Wohltmann, I., Zanis, P., and Zerefos, C.: Ozone trends at northern mid- and high latitudes – a European perspective, *Ann. Geophys.*, 26, 1207–1220, doi:10.5194/angeo-26-1207-2008, 2008.
- Harville, D.: Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems, *J. Amer. Statist. Assoc.*, 72, 320–338, 1977.
- Hastie, T. J. and Tibshirani, R. J.: *Generalized Additive Models*, Chapman&Hall, London, 1990.
- Holton, J.: The influence of gravity-wave breaking on the general-circulation of the middle atmosphere, *J. Atmospheric Sci.*, 40, 2497–2507, 1983.
- Hood, L. L. and Soukharev, B. E.: Interannual Variations of Total Ozone at Northern Midlatitudes Correlated with Stratospheric EP Flux and Potential Vorticity, *J. Geophys. Res.*, 62, 3724–3740, 2005.
- Hood, L. L., Rossi, S., and Beulen, M.: Trends in lower stratospheric zonal winds, Rossby wave breaking behavior, and column ozone at northern mid-latitudes, *J. Geophys. Res.*, 104, 24321–24339, 1999.
- Hurst, D. F., Oltmans, S. J., Vömel, H., Rosenlof, K. H., Davis, S. M., Ray, E. A., Hall, E. G., and Jordan, A. F.: Stratospheric water vapor trends over Boulder, Colorado: Analysis of the 30 year Boulder record, *J. Geophys. Res.*, 116, D02306, doi:10.1029/2010JD015065, 2011.
- Jun, M.: Non-stationary Cross-Covariance Models for Multivariate Processes on a Globe, *Scand. J. Stat.*, 38, 726–747, 2011.
- Kobayashi, C. and Shibata, K.: Evaluation of Dynamical Contribution to Lower Stratospheric Ozone Trends in Northern Mid-latitudes over the Last Three Decades (1980–2006) Using a Chemical Transport Model, *J. Meteorol. Soc. Jpn.*, 89, 363–376, 2011.
- Konopka, P., Groß, J.-U., Günther, G., Ploeger, F., Pommrich, R., Müller, R., and Livesey, N.: Annual cycle of ozone at and above the tropical tropopause: observations versus simulations with the Chemical Lagrangian Model of the Stratosphere (CLaMS), *Atmos. Chem. Phys.*, 10, 121–132, doi:10.5194/acp-10-121-2010, 2010.
- Krivobokova, T. and Kauermann, G.: A note on penalized spline smoothing with correlated errors, *J. Amer. Statist. Assoc.*, 102, 1328–1337, 2007.
- Laube, J. C., Keil, A., Bönisch, H., Engel, A., Röckmann, T., Volk, C. M., and Sturges, W. T.: Observation-based assessment of stratospheric fractional release, lifetimes, and ozone depletion potentials of ten important source gases, *Atmos. Chem. Phys.*, 13, 2779–2791, doi:10.5194/acp-13-2779-2013, 2013.
- Mäder, J. A., Staehelin, J., Brunner, D., Stahel, W. A., Wohltmann, I., and Peter, T.: Statistical modeling of total ozone: Selection of appropriate explanatory variables, *J. Geophys. Res.-Atmos.*, 112, D11108, doi:10.1029/2006JD007694, 2007.

- Manne, G. L., Santee, M. L., Rex, M., Livesey, N. J., Pitts, M. C., Veefkind, P., Nash, E. R., Wohltmann, I., Lehmann, R., Froidevaux, L., Poole, L. R., Schoeberl, M. R., Haffner, D. P., Davies, J., Dorokhov, V., Gernandt, H., Johnson, B., Kivi, R., Kyrö, E., Larsen, N., Levelt, P. F., Makhtas, A., McElroy, C. T., Nakajima, H., Parrondo, M. C., Tarasick, D. W., von der Gathen, P., Walker, K. A., and Zinoviev, N. S.: Unprecedented Arctic ozone loss in 2011, *Nature*, 478, 469–475, 2011.
- Marra, G. and Wood, S. N.: Practical variable selection for generalized additive models, *Comput. Stat. Data. An.*, 55, 2372–2387, 2011.
- Marra, G. and Wood, N. W.: Coverage properties of confidence intervals for generalized additive model components, *Scand. J. Stat.*, 39, 53–74, 2012.
- Meiring, W.: Oscillations and Time Trends in Stratospheric Ozone Levels: A Functional Data Analysis Approach, *J. Am. Stat. Assoc.*, 102, 788–802, 2007.
- Miller, A., Cai, A., Tiao, G., Wuebbles, D., Flynn, L., Yang, S., Weatherhead, E., Fioletov, V., Petropavlovskikh, I., Meng, X., Guillas, S., Nagatani, R., and Reinsel, G.: Examination of ozonesonde data for trends and trend changes incorporating solar and arctic oscillation signals, *J. Geophys. Res.*, 111, D13305, doi:10.1029/2005JD006684, 2006.
- Molina, M. J. and Rowland, F. S.: Stratospheric sink for chlorofluoromethanes - chlorine atomic-catalysed destruction of ozone, *Nature*, 249, 810–812, 1974.
- Nair, P. J., Godin-Beekmann, S., Kuttippurath, J., Ancellet, G., Goutail, F., Pazmiño, A., Froidevaux, L., Zawodny, J. M., Evans, R. D., Wang, H. J., Anderson, J., and Pastel, M.: Ozone trends derived from the total column and vertical profiles at a northern mid-latitude station, *Atmos. Chem. Phys.*, 13, 10373–10384, doi:10.5194/acp-13-10373-2013, 2013.
- Newchurch, M. J., Yang, E.-S., Cunnold, D. M., Reinsel, G. C., and Zawodny, J. M.: Evidence for slowdown in stratospheric ozone loss: First stage of ozone recovery, *J. Geophys. Res.*, 108, doi:10.1029/2003JD003471, 2003.
- Nychka, D.: Bayesian confidence intervals for smoothing splines, *J. Amer. Statist. Assoc.*, 83, 1134–1143, 1988.
- Oberländer, S., Langematz, U., Matthes, K., Kunze, M., Kubin, A., Harder, J., Krivova, N. A., Solanki, S. K., Pagaran, J., and Weber, M.: The influence of spectral solar irradiance data on stratospheric heating rates during the 11 year solar cycle, *Geophys. Res. Lett.*, 39, L01801, doi:10.1029/2011GL049539, 2012.
- Opsomer, J., Wang, Y., and Yang, Y.: Nonparametric Regression With Correlated Errors, *Stat. Sci.*, 16, 134–153, 2001.
- Petropavlovskikh, I., Bhartia, P. K., and DeLuisi, J.: New Umkehr ozone profile retrieval algorithm optimized for climatological studies, *Geophys. Res. Lett.*, 32, L06805, doi:10.1029/2005GL023323, 2005.
- Petropavlovskikh, I., Evans, R., McConville, G., Oltmans, S., Quincy, D., Lantz, K., Disterhoft, P., Stanek, M., and Flynn, L.: Sensitivity of Dobson and Brewer Umkehr ozone profile retrievals to ozone cross-sections and stray light effects, *Atmos. Meas. Tech.*, 4, 1841–1853, doi:10.5194/amt-4-1841-2011, 2011.
- Pinheiro, J. and Bates, D.: *Mixed-Effects Models in S and S-Plus*, Springer, New York, 2000.
- Ramsay, J. and Silverman, B.: *Functional Data Analysis*, Springer, New York, 2005.
- Randel, W. J., Wu, F., Vömel, H., Nedoluha, G. E., and Forster, P.: Decreases in stratospheric water vapor after 2001: Links to changes in the tropical tropopause and the Brewer-Dobson circulation, *J. Geophys. Res.*, 111, D12312, doi:10.1029/2005JD006744, 2006.
- Reinsel, G. C.: Trend analysis of upper stratospheric Umkehr ozone data for evidence of turnaround, *Geophys. Res. Lett.*, 29, doi:10.1029/2002GL014716, 2002.
- Rieder, H. E., Staehelin, J., Maeder, J. A., Peter, T., Ribatet, M., Davison, A. C., Stübi, R., Weihs, P., and Holawe, F.: Extreme events in total ozone over Arosa – Part 2: Fingerprints of atmospheric dynamics and chemistry and effects on mean values and long-term changes, *Atmos. Chem. Phys.*, 10, 10033–10045, doi:10.5194/acp-10-10033-2010, 2010.
- Rieder, H. E., Frossard, L., Ribatet, M., Staehelin, J., Maeder, J. A., Di Rocco, S., Davison, A. C., Peter, T., Weihs, P., and Holawe, F.: On the relationship between total ozone and atmospheric dynamics and chemistry at mid-latitudes – Part 2: The effects of the El Niño/Southern Oscillation, volcanic eruptions and contributions of atmospheric dynamics and chemistry to long-term total ozone changes, *Atmos. Chem. Phys.*, 13, 165–179, doi:10.5194/acp-13-165-2013, 2013.
- Robinson, G. K.: That BLUP is a good thing; the estimation of random effects (with discussion), *Stat. Sci.*, 6, 15–51, 1991.
- Ruppert, D., Wand, M. P., and Carroll, R. J.: *Semiparametric regression*, Cambridge University Press, Cambridge, 2003.
- Seppala, A., Verronen, P., Kyrola, E., Hassinen, S., Backman, L., Hauchecorne, A., Bertaux, J., and Fussen, D.: Solar proton events of October–November 2003: Ozone depletion in the Northern Hemisphere polar winter as seen by GOMOS/Envisat, *Geophys. Res. Lett.*, 31, L19107, doi:10.1029/2004GL021042, 2004.
- Silverman, B. W.: Some aspects of the spline smoothing approach to non-parametric regression curve fitting, *J. Roy. Stat. Soc B*, 47, 1–52, 1985.
- Silverman, B. W.: Smoothed functional principal components analysis by choice of norm, *Ann. Stat.*, 24, 1–24, 1996.
- Sofieva, V. F., Kalakoski, N., Verronen, P. T., Päiväranta, S.-M., Kyrölä, E., Backman, L., and Tamminen, J.: Polar-night O₃, NO₂ and NO₃ distributions during sudden stratospheric warmings in 2003–2008 as seen by GOMOS/Envisat, *Atmos. Chem. Phys.*, 12, 1051–1066, doi:10.5194/acp-12-1051-2012, 2012.
- Solomon, S.: Stratospheric ozone depletion: A review of concepts and history, *Rev. Geophys.*, 37, 275–316, 1999.
- Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davis, S. M., Sanford, T. J., and Plattner, G.-K.: Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming, *Science*, 327, 1219–1223, doi:10.1126/science.1182488, 2010.
- Steinbrecht, W., Köhler, U., Claude, H., Weber, M., Burrows, J. P., and van der A, R. J.: very high ozone columns at northern mid-latitudes in 2010, *J. Geophys. Res.*, 38, L06803, doi:10.1029/2010GL046634, 2011.
- Wahba, G.: Bayesian confidence intervals for the cross validated smoothing spline, *J. Roy. Stat. Soc B*, 45, 133–150, 1983.
- Wahba, G.: *Spline models for observational data*, Philadelphia: SIAM, 1990.
- Wahba, G., Wang, Y., Gu, C., Klein, R., and Klein, B.: Smoothing Spline ANOVA for Exponential families, with application to the

- Wisconsin Epidemiological study of diabetic retinopathy, *Ann. Stat.*, 23, 1865–1895, 1995.
- Weiss, A. K., Staehelin, J., Appenzeller, C., and Harris, N. R. P.: Chemical and dynamical contributions to ozone profile trends of the Payerne (Switzerland) balloon soundings, *J. Geophys. Res.*, 106, 685–694, 2001.
- WMO: Scientific Assessment of Ozone Depletion: 2006, Tech. rep., Global Ozone Research and Monitoring Project, Geneva, 2007.
- WMO: Scientific Assessment of Ozone Depletion: 2010, Tech. Rep. 52, World Meteorological Organisation, Global Ozone Res. and Monit. Proj., Geneva, Switzerland, 2011.
- Wohltmann, I., Lehmann, R., Rex, M., Brunner, D., and Mader, J. A.: A process-oriented regression model for column ozone, *J. Geophys. Res.-Atmos.*, 112, D12304, doi:10.1029/2006JD007573, 2007.
- Wood, S.: *Generalized Additive Models An introduction with R*, Chapman and Hall CRC Press, Boca Raton, 2006.