Atmos. Chem. Phys. Discuss., 13, 12337–12387, 2013
www.atmos-chem-phys-discuss.net/13/12337/2013/
doi:10.5194/acpd-13-12337-2013

Atmospheric Chemistry and Physics Discussions

Trends in stratospheric ozone profiles using functional mixed models

A. Y. Park[1], **S. Guillas**[1], **and I. Petropavlovskikh**[2]

[1]Department of Statistical Science, University College London, London, UK
[2]CIRES, University of Colorado, Boulder, Colorado, USA

Received: 17 April 2013 – Accepted: 25 April 2013 – Published: 14 May 2013

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

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

This paper is devoted to the modeling of altitude-dependent patterns of ozone variations over time. Umkher 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 as it penalizes roughness of the function and smooths excessive variations in the shape of the ozone profiles. As a result, data driven basis functions are obtained. Secondly we estimate the effects of covariates – month, year (trend), quasi biennial oscillation, the Solar cycle, arctic oscillation and the El Niño/Southern Oscillation cycle – on the principal component scores of ozone profiles over time using generalized additive models. The effects are smooth functions of the covariates, and are represented by knot-based regression cubic splines. Finally we employ generalized additive mixed effects models incorporating a more complex error structure that reflects the observed seasonality in the data. The analysis provides more accurate estimates of influences and trends, together with enhanced uncertainty quantification. 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. The strongly declining trends over 2003–2011 for altitudes of 32–64 hPa show that stratospheric ozone is not yet fully recovering.

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 (WMO, 2007, 2011). The discovery of the Antarctic ozone hole in early 1980s,

ACPD
13, 12337–12387, 2013

see e.g. Farman et al. (1985) and Solomon (1999), was very recently followed by the discovery of a new ozone hole in the Arctic observed for 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 in production of ChloroFluoroCarbons (CFCs), but rather due to an increase in Green-House Gases (GHGs), which warm troposphere, but increase stratospheric cooling that in turn may slow ozone depletion. Chemistry-climate models do not yet well simulate these interactions, or do it with large uncertainties, and some joint effort by the CCMVal and CCMVal-2 projects focuses on intercomparisons of such models, see e.g. 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.

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 each of the 15 Brewer altitude bands, 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 has been 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 profiles in a single model approach. Meiring (2007) smoothed vectors of ozone profiles at one location

ACPD
13, 12337–12387, 2013

(Hohenpeissenberg), carried out Principal Component Analysis (PCA) of these vectors, and smoothed again the resulting principal components. This enabled the representation of patterns of variation across altitudes to identify major sources of variation. The scores 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 data analysis. Instead, the evidence of the Solar cycle was obtained through the estimated time-dependent effect curves which exhibited peaks in 1970, 1981 and 1992, at the times the Solar cycle was at its maximum. In Meiring (2007), the model is separately fitted for month and 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 model that includes month, year, QBO, the Solar cycle, AO and El Niño/Southern Oscillation (ENSO) cycle as additive terms. Similar to Meiring (2007), we carry out an initial smoothing step. However, we enhance the principal component decomposition through a combined functional principal component decomposition that integrates two steps of data smoothing prior and after the PCA. Furthermore, by modeling all months in one setting, we acquire information that is present across months, so the fitted curves of covariates are easier to interpret as seasonal effects are already included in analysis. By including the Solar cycle and AO, we separate the effect of these two factors from yearly trends to obtain more precise trend estimates that correspond more genuinely to variations due to changing emissions of Ozone Depleting Substances (ODS) and GHGs. We also tried to model the interactions between covariates but the effects were negligible thus only additive terms were kept. Autocorrelation in the residuals was also found to be negligible. Finally, we employ Generalized Additive Mixed effects Models (GAMMs) accommodating observed heteroskedasticity in the residuals that reflect unexplained variations that are not purely noise. GAMMs

ACPD
13, 12337–12387, 2013

have two advantages over SSANOVA models. First, a complex error structure can be incorporated in the fitting process. Secondly, when the correlation or heteroskedasticity are misspecified, smoothing parameters obtained from a mixed model framework are often more reliable. As a result, we obtain trends that depict the influence of the yearly evolution of emissions on the five most important variations in the profiles.

The paper is structured as follows. In Sect. 2 we describe the representation of the ozone profiles as functional data. Section 3 explains the decomposition of these functional data using Functional Principal Component Analysis (FPCA) in order to perform dimension reduction. FPCA captures the primary modes of variation with a very parsimonious model. Section 4 focuses on the use of the Generalized Additive Models (GAMs), whereas Sect. 5 underscores the benefit of employing the GAMMs instead of the GAMs in our setting. Section 6 deals with conclusions and further discussion.

2 Data processing

2.1 Data description

Umkehr daily ozone observations from January 1978 to December 2011 in Arosa and Boulder of latitude 46° 47′ 0″ N and 40° 0′ 54″ N, respectively, are used. Ozone profiles are retrieved in sub-layers (where width is defined in log pressure scale, such as change in pressure between the top and bottom is 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 archival (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 60. It fully covers the troposphere and the stratosphere and partially covers the mesosphere. If a station is located above the sea-level, the

ACPD
13, 12337–12387, 2013

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. Hence, we focus on layers 29–60, corresponding to roughly 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 therefore, data are unequally spaced in time; thus monthly data are created by averaging out daily record. Months for which no profile was observed created missing monthly profiles, e.g. for Boulder 11 monthly profiles were missing in years 1978, 1979, 1982, 1983, 1998, 2003 and 2005 and for Arosa 4 monthly profiles were missing in 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 ozone profiles were divided into discrete layers, we view them as curves, which is determined by the degree of smoothing of the Umkehr method. With a functional representation that accommodates smoothness, the values can be more realistically evaluated and understood. For each station, functional ozone values corresponding to the month i and layer j were observed:

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

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

Let us briefly introduce how the functional ozone values are approximated by a B-splines basis system. The literature that deals with functional data may use an expansion involving B-spline basis functions $\{\phi_k(x), k = 1, \ldots, K\}$ and their associated

ACPD
13, 12337–12387, 2013

coefficients $\{c_{ik}, k = 1, \ldots, K\}$ to achieve smoothing:

$$y_{ij} = y_i(x_j) + \epsilon_{ij}, \tag{2}$$
$$y_i(x_j) = \sum_{k=1}^{K} c_{ik}\phi_k(x_j),$$

where $y_i(x_j)$ is the smooth ozone profile of month i at altitude level x_j. Spline functions are piecewise polynomials, with the polynomial pieces joining together at so called knots and possesses continuity condition and a high degree of smoothness, see e.g., de Boor (2001). Each B-spline basis function $\phi_k(x)$ is defined by knot locations thus we have to choose the knot locations and the number of basis functions K in order to define the basis system. The degree of smoothness is controlled by K. If the number of knots are chosen to be equal to the number of discretized points j we achieve interpolation. The choice of the locations and the number of knots is computationally expensive. A roughness penalty approach, see e.g. Ramsay and Silverman (2005), might alleviate the heavy computational costs associated with choosing the locations and the number of knots. 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 with a smoothing parameter takes care of excessive variations resulting from a large K. We measure roughness of the function $y_i(x)$ by $\int [D^2 y_i(x)]^2 dx$, where $D^m y_i(x)$ is the mth order derivative of the function $y_i(x)$. Thus, from Eq. (2) we measure the roughness of the function $y_i(x)$ by $\sum_{k=1}^{K} \int [c_{ik} D^2 \phi_k(x)]^2 dx$.

We placed 32 knots at each layer $x_{29}, \ldots, x_{60}$, consequently in our situation $K = 35$, see Appendix A1. The coefficient vector is estimated by the following equation

$$\min_{c_{ik}} \sum_{j=1}^{32} [y_{ij} - \sum_{k=1}^{35} c_{ik}\phi_k(x_j)]^2 + \lambda \sum_{j=1}^{32} \sum_{k=1}^{35} [c_{ik} D^2 \phi_k(x_j)]^2, \tag{3}$$

where λ is a smoothing parameter. The first term quantifies goodness of fit but the second term penalizes the roughness of the function. Instead of the number K controlling

ACPD
13, 12337–12387, 2013

the degree of smoothing, in the roughness penalty approach the smoothing parameter λ determines the level of smoothing. We used `smooth.basis` in `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 values against λ did not pin down a particular value of the parameter because GCV values were almost invariant regardless of the values of λ provided that the values are approximately smaller than 10^{-5}, so we selected $\lambda = 10^{-5}$ for both Boulder and Arosa. The smoothed monthly ozone profiles are displayed in Fig. 1.

3 Functional principal component analysis

An expansion with B-spline basis functions presented in Sect. 2.2 achieves dimension reduction but the dimension ($K = 35$) is still rather large. We now consider FPCA to reduce dimensionality further, and therefore represent each ozone profile in a more parsimonious way. The Karhunen-Loéve expansion, e.g. in Bosq (2000), tells us

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

where $\bar{y}(x) = \frac{1}{n}\sum_{i=1}^{n} y_i(x)$ is the sample mean, $\xi_l(x)$ are the Principal Components (PCs) and $\theta_{il} = \int y_i(x)\xi_l(x)dx$ are the PC scores associated with ith sample and lth PC. The PCs are built as in the multivariate case, using eigenfunctions of the sample variance covariance function $\widehat{\Gamma}(x_j, x_s) = \frac{1}{n}\sum_{i=1}^{n} y_i(x_j)y_i(x_s)$ for all x_j, x_s. We used the truncated Karhunen-Loéve expansion thus only d PCs are retained and d controls the level of fit. Since the ozone profiles were observed at discretized points $(x_{29}, \ldots, x_{60})$ the estimates of the PCs were obtained as vectors, e.g. $\boldsymbol{\xi}_l = [\xi_l(x_{29}), \ldots, \xi_l(x_{60})]^T$ and let us denote the estimate of $\boldsymbol{\xi}_l$ as $\widehat{\boldsymbol{\xi}}_l$. We assume that each PC is smooth, which implies that $\xi_l(x_j)$ and $\xi_l(x_s)$ are not likely too differ if x_j is close enough to x_s. Only

ACPD
13, 12337–12387, 2013

5 PCs, which are responsible for 99.6 % and 99.4 % of variability in Boulder and Arosa respectively, were retained. The PCs are often referred to an empirical basis functions since they provide basis functions to approximate the functional variable as seen from Eq. (4) and are driven by the data. The scores of the 5 PCs will be used for further analysis. We compute the time series of scores for /th PC by

$$\theta_{li} = \widehat{\boldsymbol{\xi}}_l^T \boldsymbol{y}_i \quad (i = 1, \ldots, n), \tag{5}$$

where $\boldsymbol{y}_i = [y_i(x_{29}), \ldots, y_i(x_{60})]^T$. For Boulder $n = 337$ as there were 11 missing months and 5 yr of volcanic periods were deleted (337 = 408 – 60 – 11) whereas for Arosa $n = 344$ as there were 4 missing months and the volcanic periods were deleted (344 = 408 – 60 – 4).

The five PCs and their associated scores are displayed in Fig. 2 and 3. Figure 2 is helpful to inspect the effect of each PC on the mean ozone variation since a PC represents variation around the mean. The size of perturbations around the mean curve, shown as (+) (–) in each panel are computed by a multiple of each PC, $\delta \times \widehat{\boldsymbol{\xi}}_l$. Conventionally, standard deviations of each PC are widely used as a multiplier δ. However, we employ here the same subjective choice of $\delta = 0.02$, to inflate the size of perturbations in order to enhance the visual quality. It is useful to point that the first five PCs are almost identical for Boulder and Arosa, as seen on the panels in columns 1 and 3 of Fig. 3. Each PC shape is associated with sensitivity of ozone profile to major geophysical or chemical parameters or combinations thereof. The attribution of each parameter to ozone variability is not that easy to disentangle. The analysis in the next sections will provide insight on these variations; for instance PC 5 can be associated with the semi-annual oscillation. The previously published research suggests seasonal and vertical partitioning of natural and anthropogenic sources that may contribute to the inter annual ozone variability that is also geographically distributed (see Intro.). These external forces are summarized by PC shapes shown in Fig. 3. In the paper we study contribution of each score that capture long-term variability of various geophysical parameters to Boulder and Arosa ozone time series. The first two PC scores show a clear annual

ACPD
13, 12337–12387, 2013

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 and therefore are marked as points in Fig. 3. The beginning of the time series tend to have rather unstable measurements, possibly related to the fact that fewer measurements were available then, see the lower right panel of Fig. 5. The following statistical analysis includes regression on known external forcings.

4 Effect of covariates on ozone variationsusing GAMs

This section focuses on finding the important sources of the unrevealed patterns and variations among the monthly ozone profiles and explaining these variations in terms of relevant covariates. In the statistical analysis we regress the 5 PC scores on known external forcings, e.g. QBO, the Solar cycle, AO and ENSO. Monthly QBO is available at http://gcmd.nasa.gov/. QBO was recorded at seven levels, 70, 50, 40, 30, 20, 15 and 10 hPa. Four-months lagged QBO data was used to account for the time lag in the QBO effect from the equator to mid latitudes. FPCA was used to reduce the dimensionality of the QBO profiles. Only the first and second dominant PC scores were kept. We denote the scores of the first PC as QBO1 and the scores of the second as QBO2. Daily Solar cycle data are available at http://www.ngdc.noaa.gov. Daily records were averaged out to create monthly Solar flux (2800 MHz Series C). Monthly AO data were obtained from http://www.cpc.ncep.noaa.gov. Monthly ENSO data were obtained from http://www.esrl.noaa.gov/psd/enso/mei/table.html.

The main purpose of this stage is to partition each score into smooth trend components and random variability. The smooth components include month, year, QBO1, QBO2, the Solar cycle, AO and ENSO. We use GAMs, e.g. Wood (2006) to fit the scores. Model specifications of the time series of scores corresponding to /th PC are

ACPD
13, 12337–12387, 2013

given by:

$$\theta_{li} = c_l + g_{l1}(\text{Month}_i) + g_{l2}(\text{Year}_i) + g_{l3}(\text{QBO1}_i) + g_{l4}(\text{QBO2}_i) + g_{l5}(\text{Solar cycle}_i) + g_{l6}(\text{AO}_i) + g_{l7}(\text{ENSO}_i) + \epsilon_{li} \quad (i = 1,\ldots,n; \quad l = 1,\ldots,5) \tag{6}$$

where the ϵ_{li} are normally distributed with mean zero and variance σ_l^2, the g_{lj} are smooth functions, and c_l is the overall mean. For the estimation, each smooth function g_{lj} is replaced by a linear combination of basis functions with associated coefficients. We used cyclic cubic splines for the month term and non-cyclic cubic splines for the rest. More details about cyclic and noncyclic regression cubic splines can be found in Wahba (1990). A splines basis system is determined by the amount and location of knots. However, it is known that they do not have a large impact on the results of the model if the smooth curves are estimated by a balance between goodness of the fit and roughness of the function. Ruppert et al. (2003) suggest that $K = \min(n(X)/4, 40)$, where K is a number of basis functions and $n(X)$ is the number of the unique values of the covariate X. Moreover, the choice of basis functions 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 regression splines were also used to approximate all smooth terms but month. However, the statistical results of two models, which involve two different basis systems, were indistinguishable. We explain in Appendix A2 how we carry out the estimation.

Figure 4 shows the estimated curves g_{lj} of the model in Eq. (6) as a solid line together with their 95 % Bayesian confidence intervals as shaded areas. It addresses the linear and nonlinear dependence of ozone levels on covariates of time and other covariates of interest. The numbers on the y-axis refer to the Effective Degrees of Freedom (EDF), the number of model parameters minus the number of constraints imposed for identifiability, of each smooth term, which assess the flexibility of the fitted model. EDF is closely related to the smoothing parameters. As the smoothing parameter increases from 0 to ∞, the EDF decreases smoothly from ∞ to 1. We cannot

ACPD
13, 12337–12387, 2013

discriminate between the linear and the insignificant effect. In Sect. 5 we employ the shrinkage method that allows the discrimination. If the smoothing parameter is large then the model is less flexible and therefore the fitted curve has very few degrees of freedom. At the opposite extreme when the smoothing parameter is zero, the maximum of EDF is achieved. When EDF= 1 then the fitted curve is a straight line. The penalty term in the fitting criterion reduces EDF of the model. Discussions of the result of GAM fits will be addressed in Sect. 5.4.

5 GAMM modeling

5.1 Motivation

In the model in Eq. (6) a particular form of the variance-covariance matrix was assumed for the error term. Graphical and numerical summaries help analyze potential shortcomings of this assumption. Observations taken on adjacent months might have stronger correlation than observations taken on non-adjacent months. This potential correlation might not be completely captured by the covariates, resulting in error correlation. In order to check the possible presence of serial correlation in the error, the residuals from the fit of GAMs was graphically inspected. However, autocorrelation plots of the residuals did not suggest that the errors are correlated, so we do not consider autocorrelation here. Note that the original scores 1, 2 and 5 present strong autocorrelation having approximate values of 0.7, 0.8 and 0.5 at lag 1 in Boulder. This allows us to conclude that the GAMs eliminated autocorrelation in original data. Scores 3 and 4 have relatively weaker autocorrelation of 0.35 and 0.4, possibly because the PC 3 and PC 4 are not strongly related to time-dependent sources of ozone variation. However, heteroscedasticity is present. Indeed, considerable variability remains even after the model is fitted, as seen in the upper left and right panels in Fig. 5. It reveals that the constant distributional assumption based on errors is not appropriate. It gives an indication that an extended additive model, allowing a more complicated error system, is

ACPD
13, 12337–12387, 2013

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 month data are counted. The counts are taken by month and year, and are displayed in the bottom panels of Fig. 5. They show a clear pattern. The plot of the log variance together with the box and whisker plot of standard residuals grouped by month show that the residual variability has a strong annual cycle.

Figure 5 shows high ozone variability in winter and spring months. It is possibly due to ozone transport in upper and middle stratosphere associated with movement of jets (both polar and subtropical) close and away from the station, and also due to stratosphere-troposphere exchange. The errors in retrieved ozone in 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, and therefore the contribution of the out-of-band light becomes more significant and therefore errors can increase (Petropavlovskikh et al., 2011).

5.2 Modeling heteroskedasticity

The upper right panel in Fig. 5 shows a periodic pattern. This pattern can be modeled by the sine and cosine wave

$$\log(\mathrm{Var}(\epsilon_{lj})) = 2\delta_{l1}\sin(2\pi\tilde{m}_i) + 2\delta_{l2}\cos(2\pi\tilde{m}_i) \quad (7)$$
$$\Leftrightarrow \mathrm{Var}(\epsilon_{lj}) = \exp(2\delta_{l1}\sin(2\pi\tilde{m}_i))\exp(2\delta_{l2}\cos(2\pi\tilde{m}_i))$$

where $\tilde{m}_i = \mathrm{month}_i/12$. A log transformation converts the multiplicative variance function to the additive one. We follow Pinheiro and Bates (2000) and the convenient arbitrary choice of $2\delta_{l1}$ and $2\delta_{l1}$. To account for the heteroscedasticity, our model in Eq. (6) is replaced with a new model in which a complex error structure is assumed. In our new model the error vector $\boldsymbol{\epsilon}_l = [\epsilon_{l1}, \ldots, \epsilon_{ln}]^T$ is assumed to be $\boldsymbol{\epsilon}_l \sim N(\mathbf{0}, \sigma_l^2 \boldsymbol{\Lambda}_l)$. $\boldsymbol{\Lambda}_l \in \mathbb{R}^{n\times n}$ is a positive definite diagonal matrix with ith diagonal elements $\frac{1}{\sigma_l^2}\mathrm{Var}(\epsilon_{lj})$,

ACPD
13, 12337–12387, 2013

see Eq. (7). By using the variance function in Eq. (7) we largely reduced the number of parameter to model heteroscedasticity. In other words, instead of estimating whole diagonal elements in $\mathbf{\Lambda}_l$ we only estimate δ_{l1} and δ_{l2}. Variance function estimation was substantially studied by Davidian and Giltinan (1995) and Ruppert et al. (2003).

5.3 Mixed effects modeling

The fitted curves from GAMs, see Fig. 4, show unnecessary roughness particularly for QBO terms. It is known that smoothing parameters derived from GCV methods are heavily affected by a misspecified error structure. Thus, smoothing parameters are often estimated within a mixed effects model framework. With the `mgcv` library in R, GAMMs can be easily handled so that we can adapt the model in Eq. (6) to include a stochastic structure. The R function `gamm` incorporates the linear mixed model framework introduced by Pinheiro and Bates (2000) into additive models by treating the smoothing parameters as variance components through likelihood theory. Pinheiro and Bates (2000) offers a full coverage of linear mixed modeling and Ruppert et al. (2003) provides a clear explanation of smoothing parameters as mixed model components.

In a mixed model framework the model matrix $\mathbf{B}$, see Appendix A2, can be partitioned into two parts, a fixed effects and a random effects. The former is absorbed into the unpenalized components, e.g. constant and linear terms, and the latter is absorbed into the penalized components. Accordingly, the matrix representation of the GAMM is:

$$\begin{aligned} \boldsymbol{\theta}_l &= \mathbf{B}_F \boldsymbol{b}_l + \mathbf{B}_R \boldsymbol{u}_l + \boldsymbol{\epsilon}_l, \\ \boldsymbol{u}_l &\sim N(0, \mathbf{\Omega}_{\boldsymbol{\lambda}_l}), \\ \boldsymbol{\epsilon}_l &\sim N(0, \sigma_l^2 \mathbf{\Lambda}_l), \end{aligned} \tag{8}$$

where $\boldsymbol{\theta}_l = [\theta_{l1}, \ldots, \theta_{ln}]^T$ is the response vector, and $\mathbf{B}_F$ and $\mathbf{B}_R$ are the model matrices corresponding to the fixed and random components. $\boldsymbol{b}_l$ is the vector of coefficients corresponding to fixed effects, e.g. coefficients of constant and linear terms. We denote the covariance matrix of the random effects as $\mathbf{\Omega}_{\boldsymbol{\lambda}_l}$ to emphasize the fact that it achieves

ACPD
13, 12337–12387, 2013

smoothing and it is determined by the vector of smoothing parameter $\boldsymbol{\lambda}_l = [\lambda_{l1}, \ldots, \lambda_{l6}]^T$. Since $\boldsymbol{u}_l$ is not a fixed parameter but a random vector we predict it rather than estimate it. If we know the parameters of the variance components, $\boldsymbol{\Omega}_{\boldsymbol{\lambda}_l}$ and σ_l^2, we can predict $\boldsymbol{u}_l$ using the conditional mean of $\boldsymbol{u}_l$ given the response $\boldsymbol{\theta}_l$, $E(\boldsymbol{u}_l|\boldsymbol{\theta}_l)$, and this approach is known as the Best Linear Unbiased Prediction (BLUP), see e.g. Robinson (1991). We explain in Appendix A3 how we carry out the estimation.

5.4 Estimation results

Figure 6 shows the estimated curves g_{jl} from GAMs (a) and GAMMs (b) as solid lines together with their 95 % Bayesian confidence intervals as shaded areas. The more complex model which include a complex stochastic structure (GAMMs) leads to modest (not that significant) changes in the estimated curves as seen in Fig. 6. The fitted trend curve from the GAMM is smoother than the one from GAM, which supports the fact that data driven smoothing parameter selection (e.g. GCV) occasionally overfits the data. Also, it is more realistic to assume that the ozone trend curve is smooth. Furthermore, Fig. 5 and 6 support the need of a more complex model. The Bayesian confidence intervals take the dissimilar variability among months into account, which manifests itself by a larger uncertainty of the estimated curve through February to April but a decrease through Autumn. Thus, a periodic pattern in the log variance of residuals is well adapted to the measure of the uncertainty. Besides, the fact that the GAMMs treat the noisy part of a curve as random might lead to more robust estimates, reducing bias and avoiding overfitting. For example, for the estimated curves representing the year effect in Fig. 6, the one in (b) is smoother than the one in (a).

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 performance are assessed across-the-function, provided that heavy over smoothing is avoided (Marra and Wood, 2012). However, intervals for smooth components that are in the penalty null space,

ACPD
13, 12337–12387, 2013

for instance straight lines, are problematic. When combined with the identifiability constraints necessary for the GAM estimation, estimates in the penalty null space often have confidence intervals that are almost of zero width at some point as seen in Fig. 6. Zero or narrow width implies that the estimation bias exceeds its variance. Consequently it undermines the theoretical argument that a Bayesian type of intervals achieve close to nominal coverage. For more discussions of confidence interval performance in GAM, see Marra and Wood (2012).

The estimated smooth curves g_{lj} from the GAMMs are displayed in Fig. 8 and 9. Variables are selected by pursuing a shrinkage approach. In order to implement model selection we replace zero values of the penalty matrix $\mathbf{S}_l$ by a small value e. The e was chosen to be very small in order to not affect the regression coefficients, except those in the penalty null space, e.g. constant and linear terms. As a result, coefficients of the smooth term (cubic spline terms) are shrunk to zero if their associated smoothing parameter is large enough. In other words, when the EDF is zero then the effect of its associated covariate is found to be 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 GAMs and provide guidance of its implementation for `mgcv` users. The fitted month curve of score 5 shows very interesting features. Indeed, according to Garcia et al. (1997), the semiannual oscillation of stratospheric ozone has an M shape that peaks at March and October and is mostly occurring in the upper stratosphere. This is consistent with the variation for altitudes corresponding to score 5 in Fig. 2 and 3. To be able to pick up such a small variation in stratospheric ozone levels is an achievement of well-tuned GAMMs.

Figure 10 depicts the estimated trends using both GAMs and GAMMs in Boulder. Nevertheless, looking at these trends, we cannot claim here that a recovery in stratospheric ozone is occurring for specific modes but we have increased our confidence that we can eventually pin them down with more time points (to improve estimates) when such changes become significant.

ACPD
13, 12337–12387, 2013

Comparing the residual plots from the model with and without random effects is appropriate for model checking. Figure 7a gives normalized residuals when constant variance is assumed and Fig. 7b demonstrates the same plot when heteroscedasticity is accounted for. In Fig. 7b the residuals are evenly distributed around zero compared to Fig. 7a where they are not. This supports the need of a more complex error structure in the model and the appropriateness of the variance function we proposed. The heteroskedastic behavior is moderated. Note that in Fig. 7a the normalized residuals are obtained by $\gamma_i = (\theta_{li} - \widehat{\theta}_{li})/\widehat{\sigma}_l$; whereas in Fig. 7b they are defined as

$$\gamma_i = \frac{\theta_{li} - \widehat{\theta}_{li}}{\widehat{\sigma}_l \sqrt{\exp(2\delta_{l1}\sin(2\pi \times \widetilde{m}_i))\exp(2\delta_{l2}\cos(2\pi \times \widetilde{m}_i))}}. \quad (9)$$

Numerical summaries of the GAMM model are presented in Table 1, 2 and 3. Table 1 indicated the level of non-linearity of the influences of the covariates. As expected, the geophysical covariates tend to be more linearly related to ozone levels than months (annual cycle) or years (highlighting a trend change), see e.g. Miller et al. (2006). Explained deviances D_l in Table 3 are defined as

$$D_l = 1 - \frac{\sum \widehat{\varepsilon}_{li}^2/(n - \mathrm{d}f)}{\sum(\theta_{li} - \bar{\theta}_{li})^2/(n-1)}, \quad (10)$$

where df is the sum of EDF of all smooth terms. 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, or significant trends, nor by the main medium-term variations of the atmosphere encapsulated considered here. Changes in dynamical parameters in the future will be associated with climate change related processes. Therefore, it is important to study these contributions and estimate their contribution to ozone variability.

ACPD
13, 12337–12387, 2013

Figure 11 displays the diagnostic plots of score1 of the GAMM. Normalized residuals obtained by Eq. (9) are examined using different graphical methods. There is a clear improvement upon the use of GAMs (see Fig. 5) as periodic patterns in residuals have disappeared. Large absolute values of residuals tend to occur in the early years of our consideration when the measurements are not considered to be reliable. Indeed, the beginning of the Boulder record has fewer measurements and the observations were done in manual mode (thus prone to operators errors). For Arosa, observational methods were established back in 1960s and were not changed until the 80s. Also a smaller number of measurements per Umkehr curve was collected prior to the automation of the measurements system in 1986. Furthermore, the abrupt changes in the size of the normalized residuals in score 1 of Boulder data are found in 2001, see the upper right panel of Fig. 11a. The score 1 capture atmospheric variability between 250 and 68 hPa. This feature could be related to the mechanism that caused abrupt decrease in the stratospheric water vapor observed over tropics in 2001 (Randel et al., 2006; Solomon et al., 2010). The changes were related to an increase in tropical upwelling and change in the transport patterns. Decrease in the low stratospheric water vapor was also observed at around the same time over Boulder (Hurst et al., 2011). Moreover, the score 1 analysis in Boulder and Arosa data displays a similar pattern (Fig. 11), with the a visible break around 2001, which makes this feature likely related to the changes in the meridional transport in the lower and middle stratosphere. This feature needs further investigation. Finally, also in score 1 for Boulder, we might detect the influence of the ENSO cycle (supported by the fact that the EDF of ENSO is 2.7) – negative in 2000–2001 and positive in 1998 – in the residuals. Indeed, ENSO influences transport patterns.

The original PC scores and the fitted PC scores using GAMMs are presented in Figs. 12 and 13. Note that scales of score values in Figs. 12 and 13 are different. Scores 3 and 4 for Arosa that have larger variability than in the case of Boulder, see also Fig. 3, which means that Boulder and Arosa have different dynamical contributions despite the fact that the PCs of Boulder and Arosa are almost identical. The fit is good, as variations

ACPD
13, 12337–12387, 2013

are captured by our explanatory variables. Even though the greater flexibility of the GAMMs allows more complicated shapes for distributions, it does not fully capture high and sharp peaks. We believe that the random spikes not accounted for by the model correspond to a smaller scale variability that is not captured by proxies, and 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) not included in the model. Figures 14 and 15 show original ozone values (in black) and estimated ozone values from the fitted scores with GAMMs (in red). Estimated ozone values at altitude of x are computed by Eq. (4) but $\widehat{\theta}_{li}$ replacing θ_{li} where $\widehat{\theta}_{li}$ are the fitted scores from GAMMs. There is some information in the upper stratosphere (e.g. 3.9 hPa) that is not captured by the model, and therefore it is important to study these processes to find their effects on long-term changes in ozone distribution in middle latitudes and in the upper stratosphere.

From GAMMs we estimated the ozone trends for each PC score, see Figs. 8 and 9. Unfortunately, for Arosa, the very same PC decomposition as for Boulder does not yield, for score 1 and score 2, realistic yearly trends. Indeed, Arosa's time series of scores 1 (in particular, and of the greatest weight) and 2 (to some extend) is more hectic than Boulder, with wider spread as well and less signal, see Fig. 3. Larger variability present in Arosa is the main explanation for the fact that Arosa is difficult to grasp, even with GAMMs. Hence there is no possibility of picking up the yearly trend, only the yearly cycle and a few other influences that may not be right as the yearly trend is not picked up. We tried to tweak the smoothing parameters used in PC decomposition, but to no avail, so it is a property of the data set. Our statistical method requires high quality of information within the time series of profiles to obtain such trend estimates. Therefore we will only consider Boulder for trend analysis. To show ozone trends that takes all PC scores into account after other effects of covariates (month, QBO, the Solar cycle, AO

ACPD
13, 12337–12387, 2013

and ENSO) are accounted for. For a given altitude level x, we compute the trend by

$$O_t(x) = \sum_{l=1}^{5} \xi_l(x)\hat{g}_{l2} \tag{11}$$

where $O_t(x)$ is the estimated ozone trend at altitude x and year t, $\hat{g}_{l2}$ are the fitted score values (only from the trend) from GAMMs and $\xi_l(x)$ are the smoothed PCs as in Eq. (4) and are computed by Eq. (A1). Estimated ozone trend curves of Boulder with their 95 % confidence intervals at selected altitudes are reported in Fig. 16. The trends seem to show the typical decrease and a beginnning of a recovery from 1996 onwards, but from 2003 at 32–8 hPa, there is a sharp decline in ozone values. The sharp decline results from the fitted trend curves of the third PC. In other words, the fitted trend curve associated with PC 3, see Fig. 10, display a sharp decline after 2003. There is also a detected increase in the troposphere, as well as a leveling off in the upper stratosphere.

To investigate further, we examine the impact of the Solar radiation. Several papers recently published results of analysis 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 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 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 level from 2004 to

ACPD
13, 12337–12387, 2013

2007. This effect is related to the spectral variability of the cycle 23 and was not found in the previous cycles. The non-linear fit of the score 4 for Solar cycle signal, see Fig. 8d, indicates that GAMM model could not produce the best fit for the Solar cycle in ozone data at the middle stratosphere and troposphere, see Fig. 3 for altitude sensitivity of 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 created different response in middle latitude ozone.

6 Conclusions

The recovery of ozone is strongly related to the remaining concentration of Equivalent Effective Stratospheric Chlorine (EESC). Recently, several publications revised life time of the CFCs and other ODS in the stratosphere (Laube et al., 2012). The age of air terminology is used to predicts how much of the ESSC is remaining 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 stratosphere. The age of air is also 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 increase in the HydroChloroFluoroCarbon (HCFC) concentrations that are used to replace CFCs under the Montreal protocol. These chemicals are more effective at ozone destruction (destroy about 10 % more ozone for the same concentration as CFC-12), and even more effective as GHGs, so there is a great need to monitor them for climate purposes. However, for our purpose of trend analysis of the stratosphere we are concerned with the continuous increase of the HCFCs in addition to other ODS, since their smaller fractional release time and longer life times of the ODS can delay ozone recovery (Laube et al., 2012).

The use of GAMMs, with seasonal variance, for ozone profiles enables better quantification of uncertainties than previous methods. It thus gives us more confidence in

ACPD
13, 12337–12387, 2013

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 non-linear fashion, adjust for heteroskedasticity, 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 more precisely the trends. But that increase in complexity of the model might not result in a significant improvement. Nevertheless, our technique only fully worked for Boulder, as we found that the information contained in the time series of profiles was large enough to allow us to compute trends at all altitudes. A recent decline at 32–64 hPa is the only surprise there.

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. 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 fit better the scores 3 and 4. One possibility would be to use reanalysis data to see if dynamical transport information could yield a larger explanatory power for the GAMM. 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).

ACPD
13, 12337–12387, 2013

Appendix A

Statistical details

A1 FPCA

Silverman (1996) proposed the method of estimating smooth PCs by incorporating a penalty term in the orthonormality constraint. Further, following Ramsay and Silverman (2005, Ch. 8–9) we used B-spline basis functions to represent the PCs, e.g. $\boldsymbol{\xi}_l = \sum_{k=1}^{K} z_{lk}\boldsymbol{\phi}_k(x)$ where $\boldsymbol{\phi}_k(x) = [\phi_k(x_{29}), \ldots, \phi_k(x_{60})]^T$, so roughness of each PC is easily quantified by $\sum_{j=1}^{32}\sum_{k=1}^{K}[z_{lk}D^2\phi_k(x_j)]^2$. We placed the knots at each layer (32) and used cubic B-splines (3) thus $K = 32+3 = 35$. After the basis expansion of the PCs the fitting criterion for the PCs becomes

$$\max \frac{\boldsymbol{z}_l^T \boldsymbol{\Phi}^T \mathbf{Y}^T \mathbf{Y} \boldsymbol{\Phi} \boldsymbol{z}_l}{\boldsymbol{z}_l^T (\mathbf{I}_{35} + \lambda \mathbf{P}^T \mathbf{P}) \boldsymbol{z}_l} \tag{A1}$$

where $\mathbf{Y} \in \mathbb{R}^{n\times 35}$ is the data matrix consisting of y_{ij}, $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1(x)|\ldots|\boldsymbol{\phi}_{35}(x)]$ with $\boldsymbol{\phi}_k(x)$ as defined above, and $\mathbf{P}^T\mathbf{P}$ is the penalty matrix that quantifies the roughness of the PCs. We aim to find a sequence of $\{\boldsymbol{z}_l, l = 1, \ldots, 5\}$ and the smoothing parameter λ. We applied the same smoothing parameter for all PCs. We selected $\lambda = 10^{-5}$. Note that the numerator is the objective function to be maximized and the denominator is the orthonormal constraint. Once $\widehat{\boldsymbol{z}}_l$ are estimated from Eq. (A1) we obtain $\widehat{\boldsymbol{\xi}}_l = \sum_{k=1}^{K}\widehat{z}_{lk}\boldsymbol{\phi}_k(x)$. The R-function `pca.fd` in FDA library computed z_{lk} and $\widehat{\boldsymbol{\xi}}_l$.

A2 GAM

To estimate the coefficients we use the penalized least squares fitting criterion without an intercept c_l:

$$\min \|\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l\|^2 + \boldsymbol{\alpha}_l^T \mathbf{S}_l \boldsymbol{\alpha}_l. \tag{A2}$$

ACPD
13, 12337–12387, 2013

Here $\boldsymbol{\theta}_l = [\theta_{l1}, \ldots, \theta_{ln}]^T \in \mathbb{R}^n$ is lth score vector, $\mathbf{B}$ is the model matrix and $\boldsymbol{\alpha}_l \in \mathbb{R}^{K_a}$ is the vector of coefficients of lth score vector. The penalty matrix $\mathbf{S}_l$ is a block diagonal matrix with blocks $\lambda_{lj}\mathbf{A}_j$ such that

$$\mathbf{S}_l = \begin{bmatrix} \lambda_{l1}\mathbf{A}_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \lambda_{l2}\mathbf{A}_2 & \cdots \\ \vdots & \vdots & \vdots \\ \mathbf{0} & \cdots & \lambda_{l7}\mathbf{A}_7 \end{bmatrix} \tag{A3}$$

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

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

The smoothing parameters λ_{lj} play a crucial role in the fitting process. GCV was used for the parameter estimation. For more discussion about the smoothing parameter selection, see e.g. Wood (2006, Ch. 4.5). The R-function `gam` in `mgcv` library was used for fitting the model in Eq. (6).

A3 GAMM

In practice, $\boldsymbol{\Omega}_{\boldsymbol{\lambda}_l}$ and σ_l^2 are unknown parameters so we must estimate them. The estimation of $\boldsymbol{\Omega}_{\boldsymbol{\lambda}_l}$ and σ_l^2 will follow shortly. Under the Gaussian assumption and the independence of $\boldsymbol{\epsilon}_l$ and $\boldsymbol{u}_l$, we deduce $\mathrm{Cov}(\boldsymbol{\theta}_l) = \mathbf{B}_R\boldsymbol{\Omega}_{\boldsymbol{\lambda}_l}\mathbf{B}_R^T + \sigma_l^2\boldsymbol{\Lambda}_l$ and denote $\boldsymbol{\Sigma}_l := \mathrm{Cov}(\boldsymbol{\theta}_l)$. Then the BLUP of $\boldsymbol{u}_l$ is

$$\widehat{\boldsymbol{u}}_l = \widehat{\boldsymbol{\Omega}}_{\boldsymbol{\lambda}_l}\mathbf{B}_R^T\widehat{\boldsymbol{\Sigma}}_l^{-1}(\boldsymbol{\theta}_l - \mathbf{B}_F\widehat{\boldsymbol{b}}_l) \tag{A5}$$

ACPD
13, 12337–12387, 2013

where $\widehat{\boldsymbol{\Omega}}_{\boldsymbol{\lambda}_l}$, $\widehat{\boldsymbol{\Sigma}}_l$ and $\widehat{\boldsymbol{b}}_l$ are the estimates of $\boldsymbol{\Omega}_{\boldsymbol{\lambda}_l}$, $\boldsymbol{\Sigma}_l$ and $\boldsymbol{b}_l$. By incorporating the random components in Eq. (8) into the error term such that $\boldsymbol{e} = \mathbf{B}_R\boldsymbol{u}_l + \boldsymbol{\epsilon}_l$ we write the model as $\boldsymbol{\theta}_l = \mathbf{B}_F\boldsymbol{b}_l + \boldsymbol{e}$ with $\boldsymbol{e} \sim N(0, \boldsymbol{\Sigma}_l)$. Then the Maximum Likelihood (ML) estimator of $\boldsymbol{b}_l$ is $\widehat{\boldsymbol{b}}_l = (\mathbf{B}_F^T\widehat{\boldsymbol{\Sigma}}_l^{-1}\mathbf{B}_F)\mathbf{B}_F\widehat{\boldsymbol{\Sigma}}_l^{-1}\boldsymbol{\theta}_l$.

Estimation of $\boldsymbol{b}_l$ and prediction of $\boldsymbol{u}_l$ requires estimation of $\boldsymbol{\Sigma}_l$ and the commonly used algorithms for the estimation of the variance components are ML and REstricted Maximum Likelihood (REML). It is known that ML estimators 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 loglikelihood function with respect to the variance parameter, the average of the likelihood over all possible values of $\boldsymbol{b}_l$ is maximized. More specifically, the loglikelihood function of $\boldsymbol{\theta}_l$

$$\mathcal{L}(\boldsymbol{\theta}_l|\boldsymbol{b}_l, \boldsymbol{\Sigma}_l) \propto -\frac{1}{2}[\log|\boldsymbol{\Sigma}_l| + (\boldsymbol{\theta}_l - \mathbf{B}_F\boldsymbol{b}_l)^T\boldsymbol{\Sigma}_l^{-1}(\boldsymbol{\theta}_l - \mathbf{B}_F\boldsymbol{b}_l)]. \tag{A6}$$

is replaced by $\mathcal{L}_R(\boldsymbol{\theta}_l|\boldsymbol{\Sigma}_l) = \int \mathcal{L}(\boldsymbol{\theta}_l|\boldsymbol{\Sigma}_l)\mathrm{d}\boldsymbol{b}_l$. For more detailed discussion on the REML criterion, see e.g. Patterson and Thompson (1971) and Harville (1977). The closed form of the integral is obtained as

$$\mathcal{L}_R(\boldsymbol{\theta}_l|\boldsymbol{\Sigma}_l) = \mathcal{L} - \frac{1}{2}\log|\mathbf{B}_F^T\boldsymbol{\Sigma}_l^{-1}\mathbf{B}_F|. \tag{A7}$$

where $\mathcal{L}$ is defined in Eq. (A6). The restricted loglikelihood in Eq. (A7) is not analytically maximized. Thus, numerical approximation (e.g. Newton–Raphson methods) is used. The `mgcv` package in R fits $\boldsymbol{\Sigma}_l$ by a moderate number of Expectation Maximization (EM) iterations to refine starting values of the variance components followed by a Newton-Raphson optimization of Eq. (A7). The consequence of REML is a more robust estimator to the misspecification of the error structure (Opsomer et al., 2001) and it avoids overfitting (Pinheiro and Bates, 2000).

ACPD
13, 12337–12387, 2013

Let us compare GAMMs with GAMs to understand how the variance-covariance matrix $\mathbf{\Omega}_{\boldsymbol{\lambda}_l}$ achieves smoothing in the context of mixed effects modeling. In GAMMs the idea of penalization can be incorporated into the variance-covariance matrix $\mathbf{\Omega}_{\boldsymbol{\lambda}_l}$ of random vector $\boldsymbol{u}_l$ via a Bayesian approach. Some prior beliefs are imposed on $\boldsymbol{u}_l$ such that $p(\boldsymbol{u}_l) \propto \exp(-\frac{1}{2}\boldsymbol{u}_l^T \sum_{j=1}^{6} \lambda_{lj} \widetilde{\mathbf{A}}_j \boldsymbol{u}_l)$ where $\widetilde{\mathbf{A}}_j$ is a sub-matrix of $\mathbf{A}_j$ in Eq. (A3) that excludes penalty null space. Note that p refers to the density function. Or equivalently, $p(\boldsymbol{\alpha}_l) \propto \exp(-\frac{1}{2}\boldsymbol{\alpha}_l^T \sum_{j=1}^{6} \lambda_{lj} \mathbf{A}_j \boldsymbol{\alpha}_l)$ where $\boldsymbol{\alpha}_l$ is the vector of coefficients without the decomposition of fixed and random effects as in Eq. (A2). From the model representation in Eq. (8) we have the conditional distribution of $\boldsymbol{\theta}_l$ given $\boldsymbol{\alpha}_l$

$$p(\boldsymbol{\theta}_l|\boldsymbol{\alpha}_l) \propto \exp\left(-\frac{1}{2\sigma_l^2}(\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l)^T \mathbf{\Lambda}_l^{-1} (\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l)\right). \tag{A8}$$

And using Bayes rule, $p(\boldsymbol{\alpha}_l|\boldsymbol{\theta}_l) \propto p(\boldsymbol{\theta}_l|\boldsymbol{\alpha}_l)p(\boldsymbol{\alpha}_l)$, we compute the posterior density of $\boldsymbol{\alpha}_l$:

$$\begin{aligned} p(\boldsymbol{\alpha}_l|\boldsymbol{\theta}_l) &\propto \exp\left(-\frac{1}{2\sigma_l^2}(\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l)^T \mathbf{\Lambda}_l^{-1}(\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l) - \frac{1}{2}\boldsymbol{\alpha}_l^T \sum_{j=1}^{6} \lambda_{lj} \mathbf{A}_j \boldsymbol{\alpha}_l\right) \\ &\propto \exp\left(-\frac{1}{2}\left(-\frac{2}{\sigma_l^2}\boldsymbol{\alpha}_l^T \mathbf{B}^T \mathbf{\Lambda}_l^{-1} \boldsymbol{\theta}_l + \boldsymbol{\alpha}_l^T \left(\frac{1}{\sigma_l^2}\mathbf{B}^T \mathbf{\Lambda}_l^{-1} \mathbf{B} + \sum_{j=1}^{7} \lambda_{lj} \mathbf{A}_j\right) \boldsymbol{\alpha}_l\right)\right), \end{aligned} \tag{A9}$$

which might be recognized as normal distribution with mean

$$E(\widehat{\boldsymbol{\alpha}}_l) = \frac{1}{\sigma_l^2}\left(\frac{1}{\sigma_l^2}\mathbf{B}^T \mathbf{\Lambda}_l^{-1} \mathbf{B} + \sum_{j=1}^{7} \lambda_{lj} \mathbf{A}_j\right)^{-1} \mathbf{B}^T \mathbf{\Lambda}_l^{-1} \boldsymbol{\theta}_l \tag{A10}$$

ACPD
13, 12337–12387, 2013

and variance

$$\mathrm{Var}(\widehat{\boldsymbol{\alpha}}_l) = \left(\frac{1}{\sigma_l^2} \mathbf{B}^T \boldsymbol{\Lambda}_l^{-1} \mathbf{B} + \sum_{j=1}^{7} \lambda_{lj} \mathbf{A}_j \right)^{-1}. \tag{A11}$$

Considering $\boldsymbol{u}_l \sim N(0, \boldsymbol{\Omega}_{\boldsymbol{\lambda}_l})$ in Eq. (8), the density of $\boldsymbol{u}_l$ is given by $p(\boldsymbol{u}_l) \propto \exp(-\frac{1}{2}\boldsymbol{u}_l^T \boldsymbol{\Omega}_{\boldsymbol{\lambda}_l}^{-1} \boldsymbol{u}_l)$ omitting terms that are not related to $\boldsymbol{u}_l$. Comparing the prior $p(\boldsymbol{u}_l) \propto \exp(-\frac{1}{2}\boldsymbol{u}_l^T \sum_{j=1}^{7} \lambda_{lj} \widetilde{\mathbf{A}}_j \boldsymbol{u}_l)$ of $\boldsymbol{u}_l$ with the density of $\boldsymbol{u}_l$, we have $\boldsymbol{\Omega}_{\boldsymbol{\lambda}_l}^{-1} = \sum_{j=1}^{7} \lambda_{lj} \widetilde{\mathbf{A}}_j$, and since $\widetilde{\mathbf{A}}_j$ are known the estimation of the variance covariance matrix of random components $\boldsymbol{\Omega}_{\boldsymbol{\lambda}_l}$ suffices to estimate of λ_{lj}.

Considering that the mean values of coefficients from the posterior distribution are written as Eq. (A10) we can show the link between the mixed effects models and penalized regression. If $\boldsymbol{\Omega}_{\lambda_l}$, σ_l^2 and $\boldsymbol{\Lambda}_l$ are known, or we condition on their estimates, then it turns out that the vector of estimated coefficients $\widehat{\boldsymbol{\alpha}}_l$ in Eq. (A10) are the values minimizing the following penalized sum of squares, see e.g. Wood (2006)

$$\frac{1}{\sigma_l^2} (\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l)^T \boldsymbol{\Lambda}_l^{-1} (\boldsymbol{\theta}_l - \mathbf{B}\boldsymbol{\alpha}_l) + \boldsymbol{\alpha}_l^T \left(\sum_{j=1}^{7} \lambda_{lj} \mathbf{A}_j \right) \boldsymbol{\alpha}_l. \tag{A12}$$

Comparing Eq. (A12) with (A2) the first term measures goodness of fit of the model but the second term penalizes the random components (e.g. spline terms) with the penalty matrix $\sum_{j=1}^{6} \lambda_{lj} \mathbf{A}_j$. Note that in GAMs the smoothing parameters that controls the degree of smoothness of fit are selected by the data, e.g. GCV; whereas in GAMMs the smoothing parameters are selected by REML.

Even though the penalty term compromises goodness of fit with the roughness of the curve, it biases the parameter estimators. As a result confidence intervals based on a frequentist approach generally give poor results with regard to realized coverage probabilities. Alternatively a Bayesian approach, originally developed by Wahba (1983)

ACPD
13, 12337–12387, 2013

and Silverman (1985) and extended by Ruppert et al. (2003) and Wood (2006) among others, is widely used. Bayesian posterior covariance matrix for the parameters is given in Eq. (A11). Since the posterior distribution of the model parameters is known, the construction of these confidence intervals is relatively straightforward. In addition to an easy implementation, simulations shows that these Bayesian intervals have good observed frequentist coverage properties, resulting from the fact that they include both a bias and variance component (Nychka, 1988; Wahba et al., 1995).

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 to "NOAA Atmospheric Composition and Climate program" as the source of funding for this paper.

References

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. 12358

Bosq, D.: Linear Processes in Functional Spaces: Theory and Applications, Springer-Verlag, New York, Inc, 2000. 12344

Craven, P. and Wahba, G.: Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation, Numer. Math., 31, 377–403, 1979. 12344

Crutzen, P. J.: Estimates of possible variations in total ozone due to natural causes and human activities, Ambio, 3, 201–210, 1974. 12338

Davidian, M. and Giltinan, D. M.: Nonlinear Models for Repeated Measurement Data, Champman & Hall, London, 1995. 12350

de Boor, C.: A Practical Guide to Splines, revisited edition, Springer, New York, 2001. 12343

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. 12339

ACPD

13, 12337–12387, 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. 12352
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. 12339
Gu, C.: Smoothing Spline ANOVA Models, Springer, New York, 2002. 12340
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. 12358
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. 12356
Harville, D.: Maximum likelihood approaches to variance component estimation and to related problems, J. Am. Stat. Assoc., 72, 320–338, 1977. 12361
Holton, J.: The influence of gravity-wave breaking on the general-circulation of the middle atmosphere, J. Atmospheric Sci., 40, 2497–2507, doi:10.1175/1520-0469(1983)040<2497:TIOGWB>2.0.CO;2, 1983. 12355
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. 12354
Jun, M.: Non-stationary cross-covariance models for multivariate processes on a globe, Scand. J. Stat., 38, 726–747, 2011. 12358
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, doi:10.2151/jmsj.2011-405, 2011. 12358
Konopka, P., Grooß, 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. 12346
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

ACPD
13, 12337–12387, 2013

depletion potentials of ten important source gases, Atmos. Chem. Phys., 13, 2779–2791, doi:10.5194/acp-13-2779-2013, 2013. 12357
Manney, G., 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., Makshtas, A., McElroy, C. T., Nakajima, H., Concepción Parrondo, M., Tarasick, D. W., von der Gathen, P., Walker, K. A., and Zinoviev, N. S.: Unprecedented Arctic ozone loss in 2011, Nature, 478, 469–475, doi:10.1038/nature10556, 2011. 12339
Marra, G. and Wood, S. N.: Practical variable selection for generalized additive models, Comput. Stat. Data. An., 55, 2372–2387, 2011. 12352
Marra, G. and Wood, N. W.: Coverage properties of confidence intervals for generalized additive model components, Scand. J. Stat., 39, 53–74, 2012. 12351, 12352
Meiring, W.: Oscillations and time trends in stratospheric ozone levels: a functional data analysis approach, J. Am. Stat. Assoc., 102, 788–802, 2007. 12339, 12340
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. 12339, 12353
Molina, M. J. and Rowland, F. S.: Stratospheric sink for chlorofluoromethanes – chlorine atomic-catalysed destruction of ozone, Nature, 249, 810–812, 1974. 12338
Nychka, D.: Bayesian confidence intervals for smoothing splines, J. Am. Stat. Assoc., 83, 1134–1143, 1988. 12364
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. 12356
Opsomer, J., Wang, Y., and Yang, Y.: Nonparametric regression with correlated errors, Stat. Sci., 16, 134–153, 2001. 12361
Patterson, H. D. and Thompson, R.: Recovery of interblock information when block sizes are unequal, Biometrika, 58, 545–554, 1971. 12361
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. 12341

ACPD
13, 12337–12387, 2013

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. 12349

Pinheiro, J. and Bates, D.: Mixed-Effects Models in S and S-Plus, Springer Verlag New York LLC, New York, 2000. 12349, 12350, 12361

Ramsay, J. and Silverman, B.: Functional Data Analysis, Springer Science Business Media Inc, New York, 2005. 12343, 12359

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. 12354

Robinson, G. K.: That BLUP is a good thing; the estimation of random effects (with discussion), Stat. Sci., 6, 15–51, 1991. 12351

Ruppert, D., Wand, M. P., and Carroll, R. J.: Semiparametric Regression, Cambridge University Press, Cambridge, 2003. 12347, 12350, 12364

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, 31, L19107, doi:10.1029/2004GL021042, 2004. 12355

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. 12364

Silverman, B. W.: Smoothed functional principal components analysis by choice of norm, Ann. Stat., 24, 1–24, 1996. 12359

Sofieva, V. F., Kalakoski, N., Verronen, P. T., Salmi, S.-M., Kyrölä, E., Backman, L., and Tamminen, J.: Changes in chemical composition of the middle atmosphere caused by sudden stratospheric warmings as seen by GOMOS/Envisat, Atmos. Chem. Phys. Discuss., 11, 23317–23348, doi:10.5194/acpd-11-23317-2011, 2011. 12355

Solomon, S.: Stratospheric ozone depletion: a review of concepts and history, Rev. Geophys., 37, 275–316, 1999. 12339

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. 12339, 12354

ACPD
13, 12337–12387, 2013

Wahba, G.: Bayesian confidence intervals for the cross validated smoothing spline, J. Roy. Stat. Soc. B, 45, 133–150, 1983. 12363
Wahba, G.: Spline models for observational data, Philadelphia, SIAM, 1990. 12347
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. 12364
WMO: Scientific Assessment of Ozone Depletion: 2006, Tech. Rep. 50, World Meteorological Organisation, Global Ozone Res. and Monit. Proj., Geneva, Switzerland, 2007. 12338
WMO: Scientific Assessment of Ozone Depletion: 2010, Tech. Rep. 52, World Meteorological Organisation, Global Ozone Res. and Monit. Proj., Geneva, Switzerland, 2011. 12338, 12339
Wood, S.: Generalized Additive Models An introduction with R, Champman and Hall CRC Press, Boca Raton, 2006. 12346, 12360, 12361, 12363, 12364

ACPD
13, 12337–12387, 2013

Table 1. EDF of each covariate.

			Boulder							Arosa				
Scores	Mon	Year	QBO1	QBO2	Solar	AO	ENSO	Mon	Year	QBO1	QBO2	Solar	AO	ENSO
score1	8	0	0	0	0	0	2.7	7.5	1	0	0.7	0	2.8	0
score2	8.4	3.1	1.5	0.8	2.7	0	0	7.9	2.5	0	1.4	1.1	0	0
score3	7.4	4.6	2.5	0.7	1.6	3.1	1.3	4.3	2.3	0	0	0.9	0	0
score4	5	2.6	1.8	0.9	2.6	1	0.9	5	1.9	0	0	0.2	2.8	0
score5	7.8	4.5	1.7	0.6	1.5	0	0	7.6	3.7	0	0	0	0	0

ACPD
13, 12337–12387, 2013

Table 2. Estimates of variance components parameters δ_1 and δ_2 in Eq. (7).

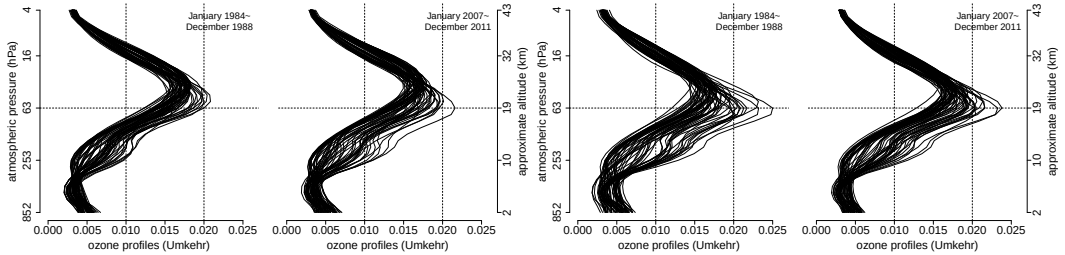
	Boulder		Arosa	
Scores	δ_1	δ_2	δ_1	δ_2
score1	0.5	0.5	0.5	0.4
score2	0.4	0.2	0.2	0.3
score3	0.4	0.4	0.3	0.3
score4	0.1	0.3	0.1	0.2
score5	0.5	0.1	0.2	0.1

ACPD
13, 12337–12387, 2013

Table 3. Explained deviance D_l (%) in Eq. (10) and the estimates of noise levels σ_l in Eq. (8).

	Boulder		Arosa	
Scores	D_l	σ_ϵ^2	D_l	σ_ϵ^2
score1	94	6.12e-06	88	1.11e-05
score2	90	1.14e-06	91	1.34e-06
score3	43	3.91e-06	31	5.84e-06
score4	40	6.03e-07	42	1.0e-06
score5	66	1.65e-07	50	3.33e-07

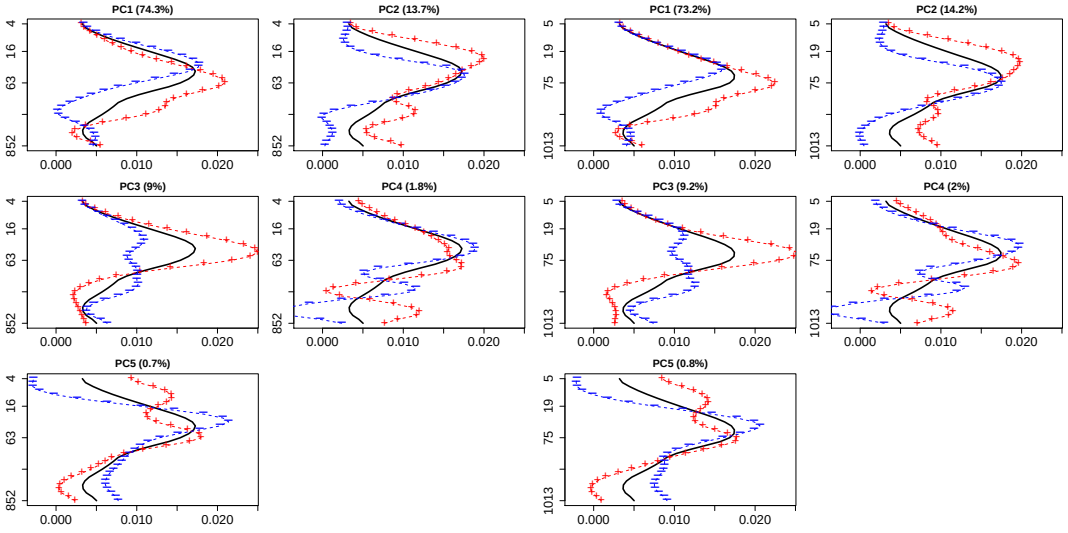
ACPD
13, 12337–12387, 2013



(a) Boulder (b) Boulder (c) Arosa (d) Arosa

Fig. 1. Smoothed monthly ozone profiles for selected periods of January 1984–December 1988 and January 2007–December 2011. In each panel, vertical lines are drawn at 0.01 and 0.02 of ozone concentrations and horizontal lines are drawn at 63 hPa, which is equivalent of 19 km in altitude. We provide the level of altitude (km) in addition to pressure level in hPa. The ozone units are in mixing ratio.

ACPD
13, 12337–12387, 2013



(a) Boulder (b) Arosa

Fig. 2. The estimated five PCs from Eq. (4). It displays the mean curve as a solid line, along with (+) and (−) indicating the 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 the ozone level and the y-axis to the atmospheric pressure (hPa). The variance contribution in % of each PC is placed on the top of each panel.

ACPD
13, 12337–12387, 2013

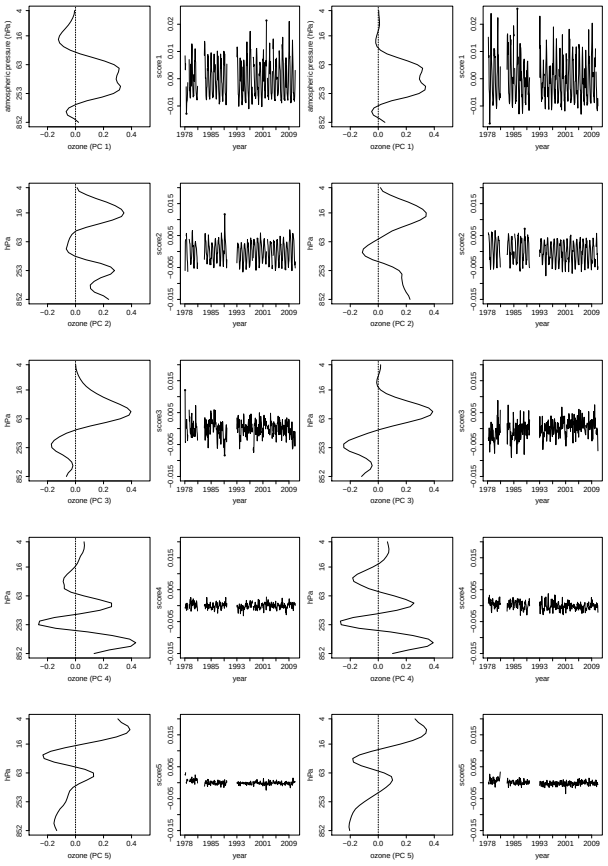
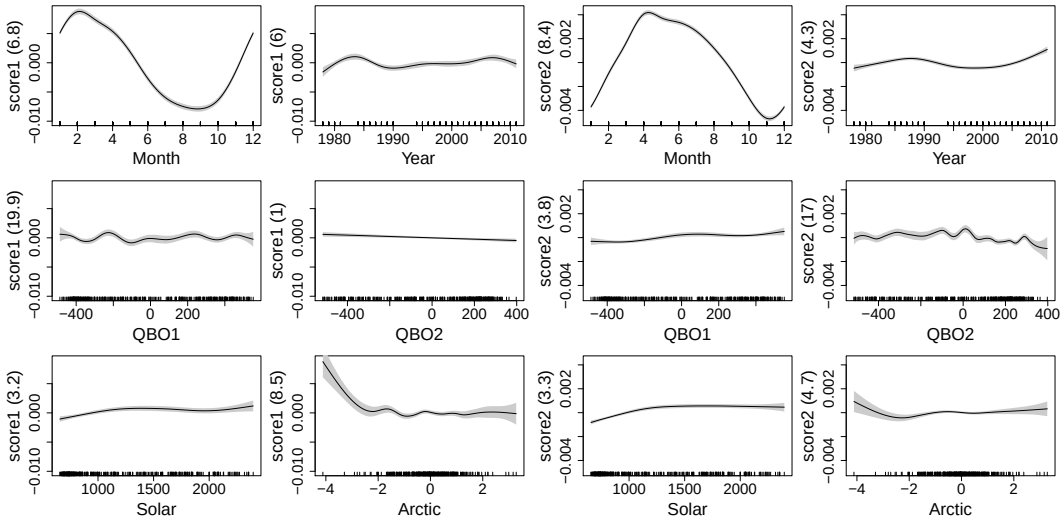


Fig. 3. The estimated five PCs, and the corresponding scores of the ozone profiles. First and third columns are the PCs for Boulder and Arosa, respectively. Second and fourth columns represent scores associated with the PCs on their left hand side. Volcanic periods (1982–1983 and 1991–1993) are omitted.

ACPD
13, 12337–12387, 2013



(a) Score1 (b) Score2

Fig. 4. The estimated smooth curves of the GAM, see Eq. (6). Only the fits of score 1 in **(a)** and 2 in **(b)** of Boulder are reported. The numbers in brackets on the y-axis captions are the estimated EDF of each smooth curve. When EDF = 1 the fitted curve is a straight line. The shaded area demonstrates 95 % Bayesian confidence intervals of each smooth curve.

ACPD
13, 12337–12387, 2013

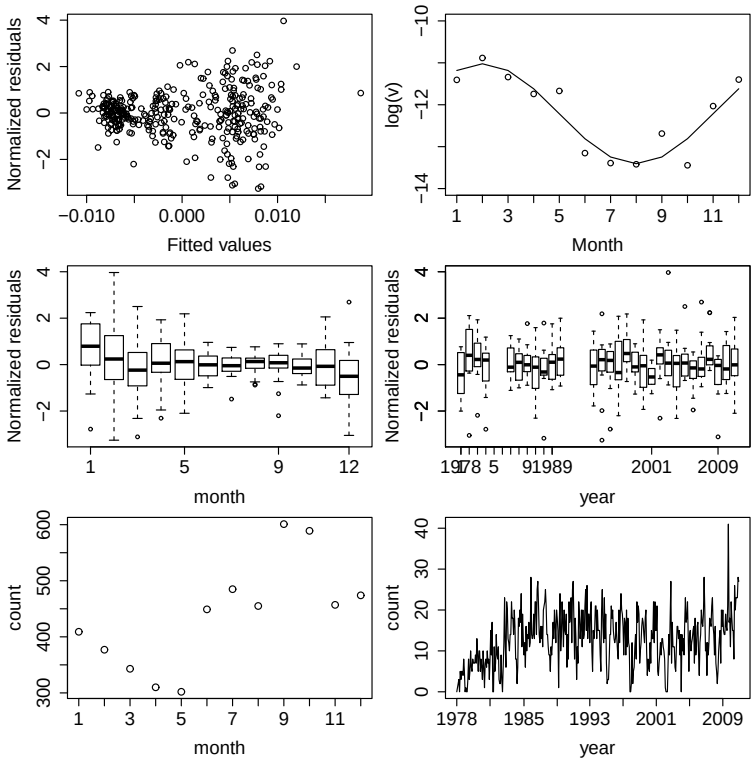
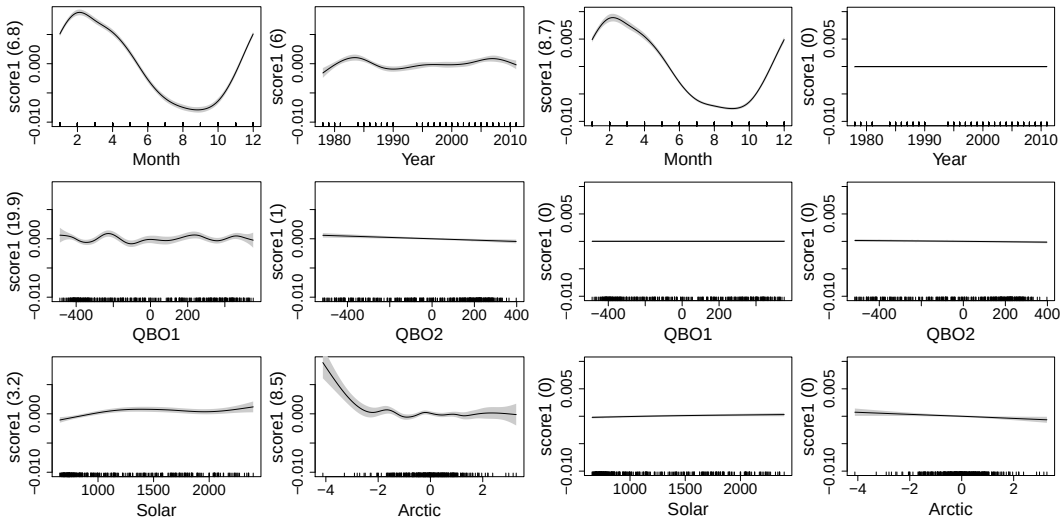


Fig. 5. Diagnostic plots of the GAM in Eq. (6) for score 1 of Boulder. Upper-left panel: plot of normalized residuals versus fitted values. Upper right panel: log variance of raw residuals computed by month. The panels in the second row are box plots of residuals by month and year. Horizontal lines are the median value. Lower 0.25 and the upper 0.75 quantiles of residuals are presented as a box. Bottom panels: counts of daily ozone observations per month and year as a time series.

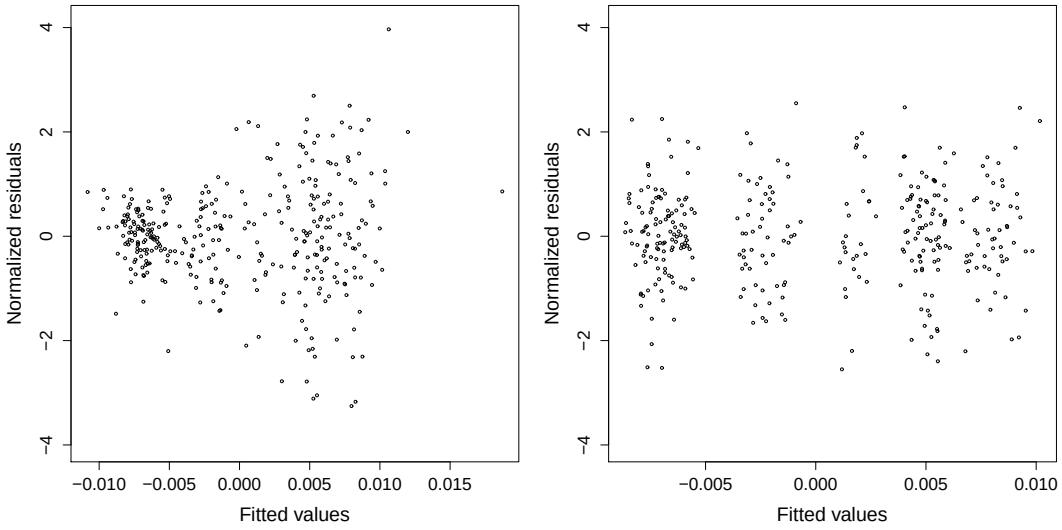
ACPD
13, 12337–12387, 2013



(a) GAMs (b) GAMMs

Fig. 6. Comparison between the results of GAM and GAMMs for score 1 of Boulder. The y-axis captions reports the EDFs of the smooth curves (EDF=1 corresponds to a straight line). Bayesian 95 % confidence bands are presented as shaded areas.

ACPD
13, 12337–12387, 2013

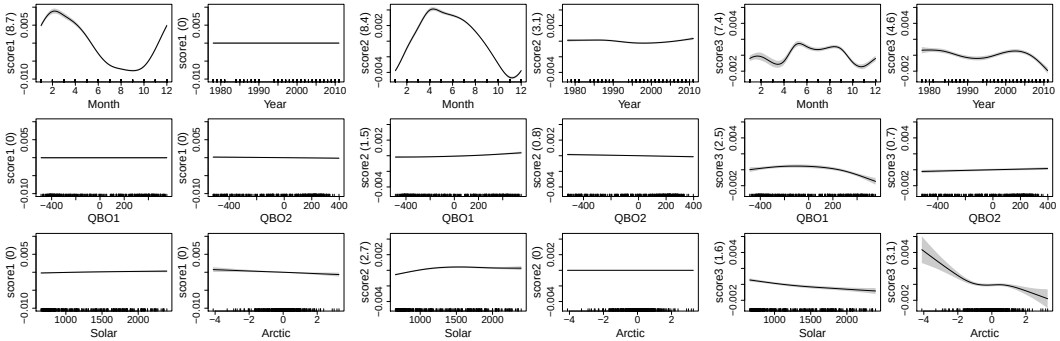


(a) GAM

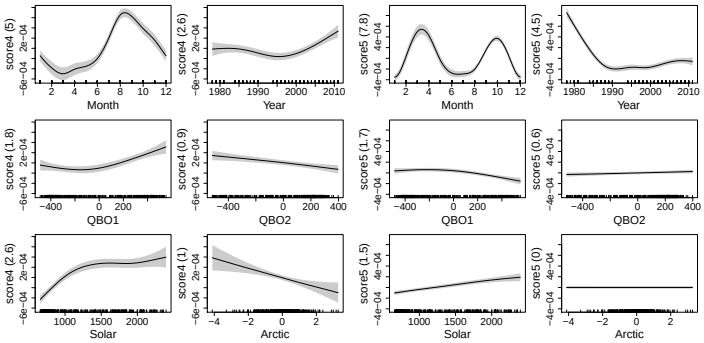
(b) GAMM

Fig. 7. Normalized residuals of the GAM in **(a)** and those of the GAMM in **(b)**. Only the residuals of score 1 of Boulder are reported.

ACPD
13, 12337–12387, 2013



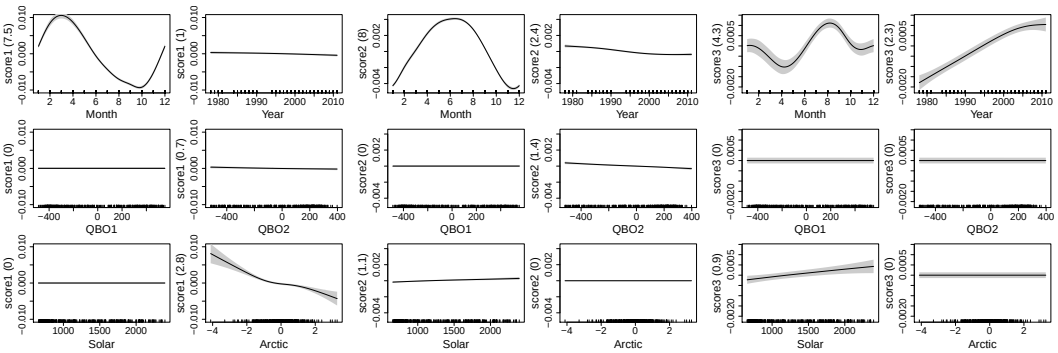
(a) score1 (b) score2 (c) score3



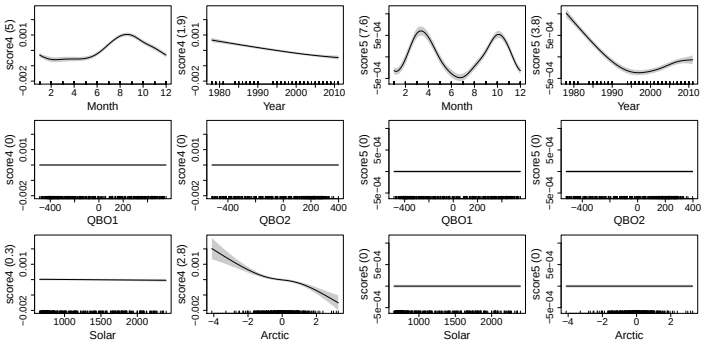
(d) score4 (e) score5

Fig. 8. Fitted smooth curves from the GAMM of Boulder. The EDFs are on the y-axis. 95 % Bayesian confidence intervals are given by shaded areas.

ACPD
13, 12337–12387, 2013



(a) score1 (b) score2 (c) score3



(d) score4 (e) score5

Fig. 9. Fitted smooth curves from the GAMMs of Arosa. The EDFs are on the y-axis. 95 % Bayesian confidence intervals are given by shaded areas.

ACPD
13, 12337–12387, 2013

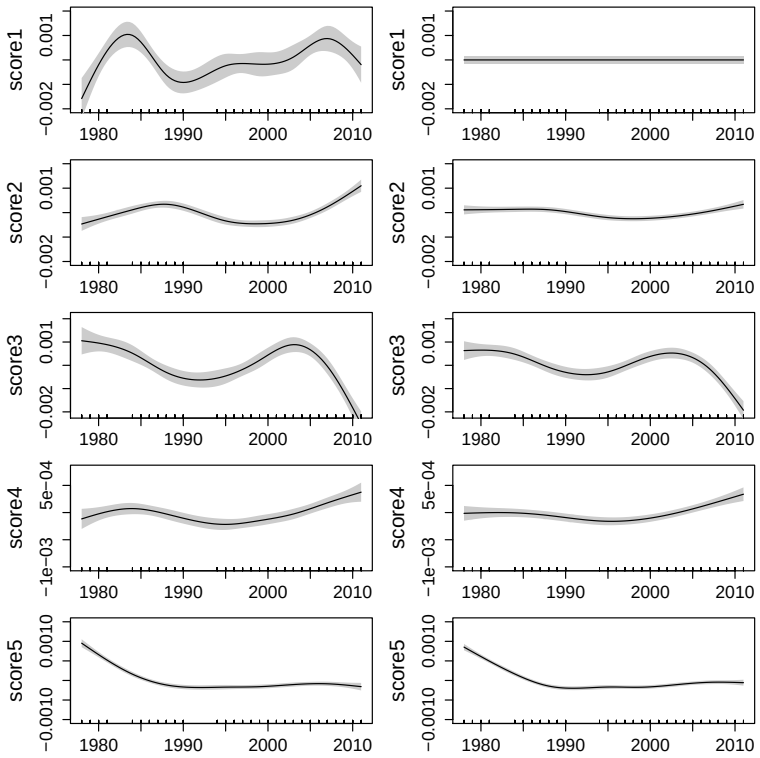
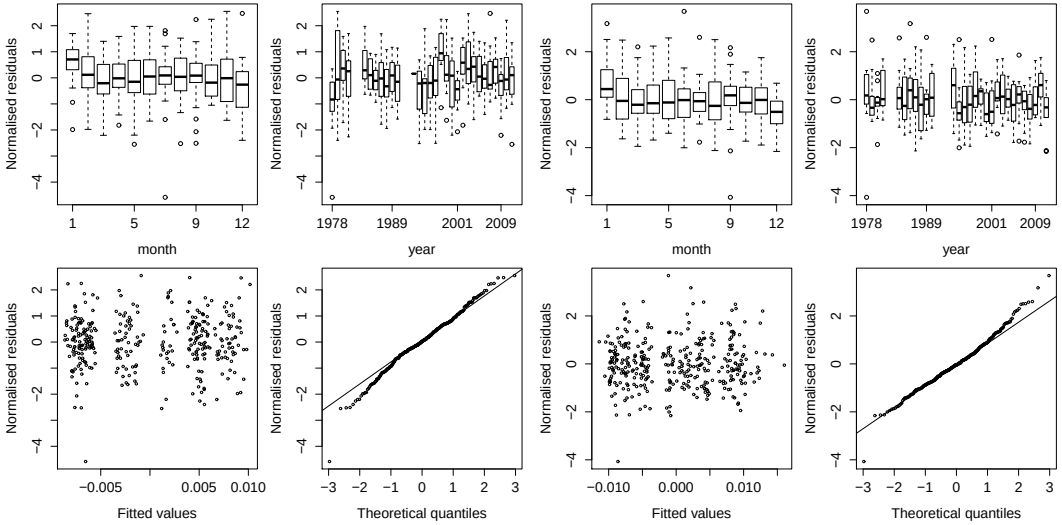


Fig. 10. Estimated trends from the GAM (left panels) and the GAMM (right panels) of Boulder for scores 1–5. Volcanic time periods removed are represented by the unticked *x* values on the x-axis. Bayesian 95 % confidence bands are represented as shaded areas.

ACPD
13, 12337–12387, 2013



(a) Boulder (b) Arosa

Fig. 11. Diagnostic plots of the GAMM for score 1. Box plots of normalised residuals against month (upper left plots of each panel), against year (upper right plots). Lower left plots of each panel: fitted values against normalized residuals. Lower right plots: Q-Q plots.

ACPD
13, 12337–12387, 2013

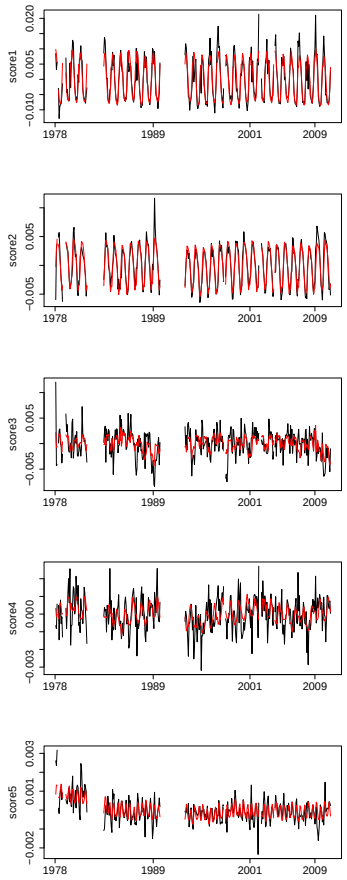


Fig. 12. Original scores (in black) and fitted scores from the GAMM (in red) of Boulder. Volcanic years are removed.

ACPD
13, 12337–12387, 2013

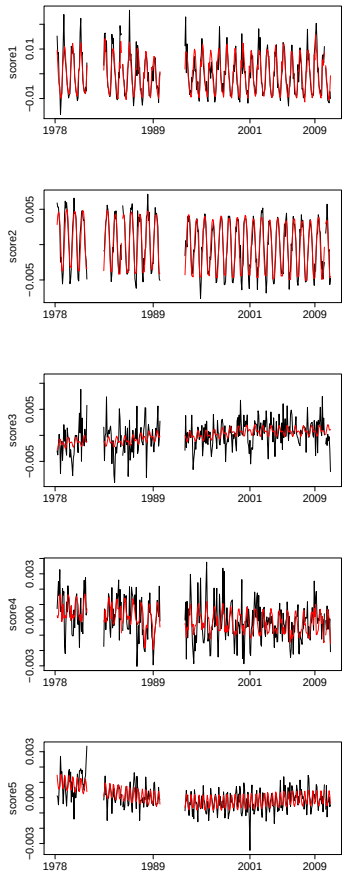


Fig. 13. Original scores (in black) and fitted scores from the GAMM (in red) of Arosa. Volcanic years are removed.

ACPD
13, 12337–12387, 2013

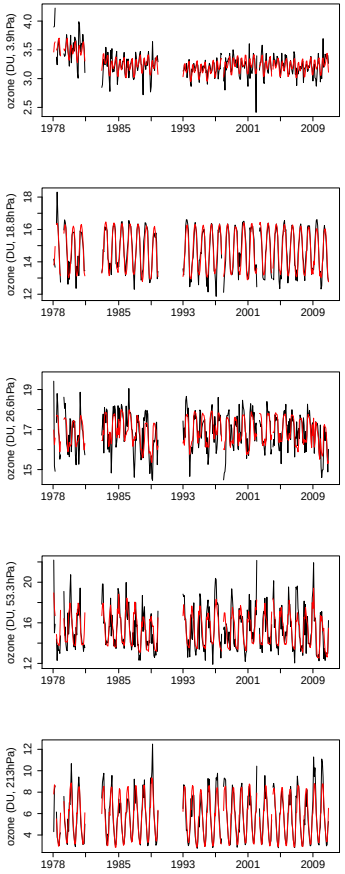


Fig. 14. Original ozone values (in black) and ozone values using fitted scores from the GAMM (in red) of Boulder. Volcanic years are removed. The ozone units are in DU.

ACPD
13, 12337–12387, 2013

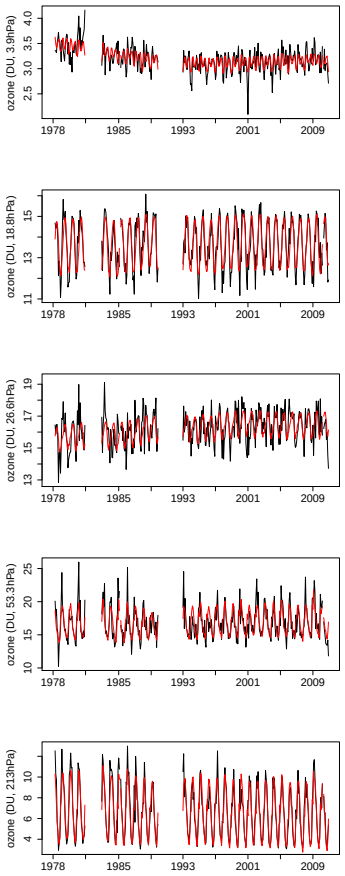


Fig. 15. Original ozone values (in black) and ozone values using fitted scores from the GAMM (in red) of Arosa. Volcanic years are removed. The ozone units are in DU.

ACPD
13, 12337–12387, 2013

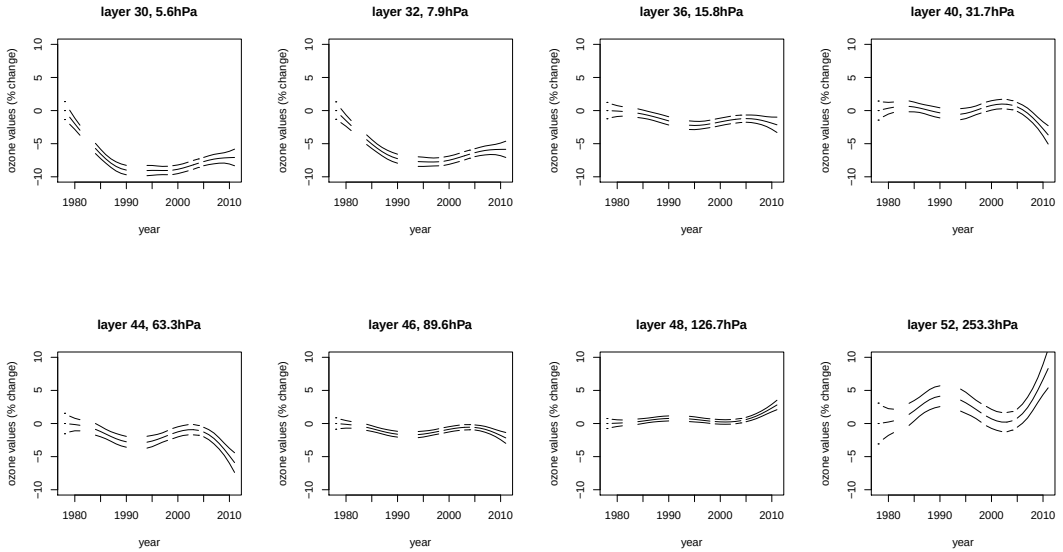


Fig. 16. Estimated ozone trend as % changes in Boulder at selected layers, with 95 % confidence intervals.

ACPD
13, 12337–12387, 2013