# Estimates of past and future ozone trends from multimodel simulations using a flexible smoothing spline methodology

John F. Scinocca,[1] David B. Stephenson,[2] Trevor C. Bailey,[2] and John Austin[3]

Received 30 November 2009; revised 19 August 2010; accepted 24 August 2010; published 24 November 2010.

[1] A novel additive model analysis of multimodel trends is presented. The approach is motivated by, and particularly suited to, the analysis of multimodel time series of varying length. This Time series Additive Model (TSAM) approach consists of three distinct steps: estimation of individual model trends, baseline adjustment of the trends, and the weighted combination of the individual model trends to produce a multimodel trend (MMT) estimate. The baseline adjustment step is not an essential ingredient of the TSAM but is included to reduce model spread. The association of the TSAM approach with a probabilistic model allows trend estimates to be used to make formal inference (e.g., calculation of confidence and prediction intervals). The method is applied to the analysis of multimodel ozone time series of varying lengths as were considered for the 2006 Scientific Assessment of Ozone Depletion. The advantages of the TSAM approach are demonstrated to include the production of smooth trend estimates out to the ends of the time series, the ability to model explicitly interannual variability about the trend estimate, and the ability to make rigorous probability statements. Calculated ozone return dates are consistent with previous qualitative estimates, but the more quantitative analysis provided by the MMT is expected to allow such data sets to be better utilized by the community and policy makers.

**Citation:** Scinocca, J. F., D. B. Stephenson, T. C. Bailey, and J. Austin (2010), Estimates of past and future ozone trends from multimodel simulations using a flexible smoothing spline methodology, *J. Geophys. Res.*, *115*, D00M12, doi:10.1029/2009JD013622.

## 1. Introduction

[2] One of the most significant advances in the 2006 Ozone assessment [*WMO*, 2007] is that future projections of ozone were being made with chemistry models embedded within upwardly extended versions of the atmospheric general circulation models used for the IPCC assessments [*Eyring et al.*, 2006, 2007]. These are often referred to as Chemistry Climate Models or CCMs. The multimodel data set used for the 2006 ozone assessment was amassed under the Chemistry-Climate Model Validation (CCMVal) Activity and is referred to as "CCMVal-1." Projections of future ozone in the CCMVal-1 data set are derived from the REF-A2 experiment, which employed the moderate (A1B) IPCC AR4 future scenario for greenhouse gas concentrations and a future scenario (B2) for the evolution of ozone depleting substances that reflects controls resulting from the Montreal Protocol and subsequent amendments and adjustments [*WMO*, 2003, Table 4B-2].

[3] While the specified period for the REF-A2 simulations spanned the period 1980–2100, due to the computational expense of the CCMs, modeling groups generally provided only a portion of the requested data. For example, individual REF-A2 contributions ranged from ensembles of one, extending over the period 2000–2019, to ensembles of three, extending over the expanded period 1960–2100. Due to the disparity in time series length and periods covered in the REF-A2 portion of the CCMVal-1 data set, the evaluation of multimodel mean time series and, therefore, multimodel estimates of ozone recovery are not straightforward. Previous trend estimation in the REF-A2 data have provided mostly qualitative results making it difficult to formulate and utilize multimodel projections [*WMO*, 2007; *Eyring et al.*, 2007].

[4] In this study we formulate a new statistical modeling approach that employs a nonparametric additive model to estimate individual-model trends (IMT), and the multimodel, trend (MMT) for time series of unequal length. Here the term "trend" refers to a smooth trajectory passing through the time series data representing the "signal" leaving a "noise" field as the residual. The goal in this procedure is the definition of the simplest nonparametric additive model whose trend estimate produces residuals that satisfy assumed properties of noise (e.g., that it be an independent normally distributed random variable). The association with a probabilistic model allows the trend estimates to be used to make formal inference (e.g., calculation of confidence and prediction intervals). We shall refer to this new approach as

[1]Canadian Centre for Climate Modeling and Analysis, Environment Canada, Victoria, British Columbia, Canada.
[2]Mathematics Research Institute, University of Exeter, Exeter, UK.
[3]University Corporation for Atmospheric Research/Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, New Jersey, USA.

the Time Series Additive Model, or "TSAM," approach. Attractive properties of the TSAM approach include: the production of smooth trend estimates out to the ends of the time series, the ability to model explicitly interannual variability about the trend estimate, and the ability to make rigorous probability statements. Because the TSAM is based on a testable probabilistic model, the suitability of the particular nonparametric additive model used can be validated.

[5] The TSAM approach adopted here consists of three steps: estimation of individual model trends (IMT), baseline adjustment of these trends, and the weighted combination of the adjusted IMT estimates to produce a multimodel trend (MMT) estimate. Much of the development effort of the TSAM approach has gone into the final weighting step. The formulation allows the specification of prior model weights if this is desired (e.g., metric-based performance weighting) in the evaluation of the final MMT estimate. Two types of uncertainty intervals are constructed for the MMT estimate. The first is the pointwise 95% confidence interval on the trend. This interval has a 95% chance of overlapping the "true trend" (i.e., the expected trend predicted by the statistical model), representing the local uncertainty in the trend at each year. The second interval, larger by construction, is the 95% prediction interval, which represents the uncertainty in predicting a value for an individual year. This interval is a combination of uncertainty in the trend estimate and uncertainty due to natural interannual variability about the trend.

[6] The present study focuses on the derivation of the TSAM approach and its basic properties. A more complete analysis of ozone in the CCMVal-1 data set as revealed by the TSAM approach is presented in the companion study [*Austin et al.*, 2010] in this special issue. *Austin et al.* [2010] also applies the TSAM approach to the newer CCMVal-2 data set [*Eyring et al.*, 2008]. Unlike CCMVal-1, essentially all model contributions to CCMVal-2 spanned the entire length of the requested period of integration (1960–2100). Application of the TSAM to both allows a quantitative comparison of, and clearer statements about, multimodel projections of ozone decline and recovery.

[7] Other studies have used observations to determine the statistical probability of detecting ozone recovery, e.g., using the method of cumulative sums [*Newchurch et al.*, 2003; *Yang et al.*, 2006; *WMO*, 2007, chapter 6]. This method is less easily applied to the situation in which a dozen or more models may be contributing. Also, *Hofmann et al.* [1997] used ozone observations over the south pole to try to detect ozone recovery. However, their analysis if extended to recent years could have resulted in quite misleading conclusions because of the 2002 southern hemisphere stratospheric warming. Those studies generally concentrated on the impact of halogen induced ozone loss which is just one part of the ozone change which occurs over a century length timescale which is the focus of attention here and in the study by *Eyring et al.* [2007]. Earlier works by *Andersen et al.* [2006] and *Weatherhead et al.* [2000] also used statistical techniques to estimate the timing of ozone recovery, but with a small number of models and simulations of short duration their conclusions would be largely superceded first by *Eyring et al.* [2007] and using the more rigorous techniques of the current paper.

[8] The outline of the paper is as follows. In section 2 the development of the TSAM approach is presented using specific time series examples from the CCMVal-1 data set. In section 3, the application of the TSAM is illustrated by using it to derive individual and multimodel estimates of ozone recovery in the form of return dates for the CCMVal-1 REF-A2 experiment. In section 4 we conclude with a brief summary and discussion.

## 2. TSAM Approach

[9] Here we introduce a statistical modeling approach that uses nonparametric regression to estimate smooth trends from time series data. The nonparametric regression uses a set of optimal thin plate splines to represent the trends and can be used to make formal inference (e.g., calculate confidence and prediction intervals). As discussed in the Introduction, the approach adopted here consists of three distinct steps: estimation of individual model trends (IMT), baseline adjustment of the trends, and the weighted combination of the adjusted individual model trends to produce a multimodel trend (MMT) estimate. In this section the development and application of this approach will be illustrated using the time series of column ozone data presented in Figure 1. These data correspond to the CCMVal-1 raw time series analyzed by *Eyring et al.* [2007, Figure 7]. Several models participating in CCMVal-1 provided two overlapping time series of column ozone to cover the maximum range of the REF-A2 period (1980–2100): one from the REF-A2 experiment and one from the climate-of-the-20th-century experiment REF-A1. This additional complication is accounted for in the TSAM approach by considering these partially overlapping time series as ensemble members.

### 2.1. Nonparametric Estimation of the Individual Model Trends

[10] The time series $y_{jk}(t)$ of an ozone-related index, such as one of those displayed in Figure 1, is additively modeled as the sum of a smooth unknown model-dependent trend, $h_j(t)$, and irregular normally distributed noise:

$$y_{jk}(t) = h_j(t) + \epsilon_{jk}(t), \tag{1}$$

where the noise field

$$\epsilon_{jk}(t) \sim N\left(0, \sigma^2\right) \tag{2}$$

is assumed to be an independent normally distributed random variable with zero mean and variance $\sigma^2$, and the indices $j$ and $k$, respectively, represent model and ensemble-member number. (Here the ensemble index $k$ extends over both REF-A1 and REF-A2 simulations for some models.) This is a nonparametric regression of the time series on time. The regression is nonparametric because the function of time does not have a fixed functional form with explicit parameters. The noise term (2), representing natural variability about the trend, is considered to be an independent normally distributed random variable: independent between different times, models, and runs. The variance of the noise is assumed to be constant over all models and runs. By fitting the trend to all the data rather than to each model

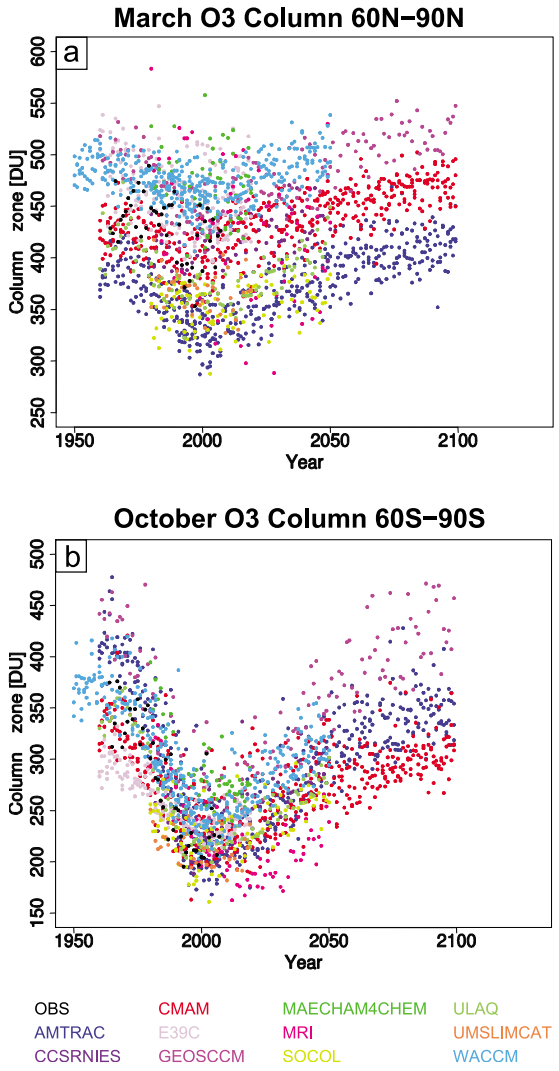


**Figure 1.** CCMVal-1 time series of monthly averaged total column ozone (a) in the latitude band 60°N–90°N for March and (b) in the latitude band 60°S–90°S for October. Following *Eyring et al.* [2007], these time series include REF-A1 data in addition to REF-A2 data for several of the models.

separately one can obtain better estimates of the noise variance (referred to as "borrowing strength").

[11] The unknown smooth functions $h_j(t)$ are estimated by fitting the data to a finite set of smooth basis functions having optimal interpolating properties. This was done here by using the generalized additive model **gam**() function in the **mgcv** library of the R language [*R Development Core Team*, 2008; *Woods*, 2006]. The default option was used which fits the data to a set of thin plate regression splines, by maximizing penalized likelihood to find the coefficients multiplying the basis functions. The smoothness of the basis functions is controlled by a smoothing parameter, which is chosen using a leave-one-out generalized cross-validation prediction approach (see *Woods* [2006] for more details). Unlike iterated 1:2:1 Lanczos filter smoothing, typically used on CCMVal-1 data [e.g., *Eyring et al.*, 2007], the thin plate splines are guaranteed to give smooth trend estimates and do not alter their properties at the ends of the series.

[12] The first step in the TSAM approach is to apply the nonparametric regression (1) to the raw time series data. This is illustrated in Figures 2a and 2b by the IMT estimates $h_j(t)$ of the CCMVal-1 March 60°N–90°N and October 60°S–90°S total column ozone displayed in Figure 1. (Note that, while the smooth trend estimates $h_j(t)$ extend over the full period (1950–2100) in Figures 2a and 2b, we have elected to display the $h_j(t)$ only over the period where data exists for each model.)

### 2.2. Baseline Adjustment of the Trend Estimates

[13] The initial IMT estimates $h_j(t)$ in Figures 2a and 2b reveal significant differences in the background values of column ozone, particularly in the Arctic (Figure 2a). To facilitate a comparison of the trends across models, anomaly time series are constructed relative to a pre–ozone-hole baseline value of the index. While this is analogous to the procedure employed by *Eyring et al.* [2007], the smoothness of $h_j(t)$ allows a more robust definition of the baseline at a particular time $t_0$ (i.e., $h_j(t_0)$), rather than from the average over some period about $t_0$. This results in the anomaly time series:

$$y_{jk}(t) - h_j(t_0). \tag{3}$$

By construction, the anomaly time series (3) is centered on a baseline value of zero at the time $t_0$. Here we chose to have this baseline changed from zero to the multimodel mean of $h_j(t_0)$ resulting in the "$t_0$ baseline-adjusted time series":

$$y'_{jk}(t) = y_{jk} - h_j(t_0) + h(t_0). \tag{4}$$

where

$$h(t_0) = \frac{1}{J}\sum_{j'=1}^{J} h_{j'}(t_0), \tag{5}$$

where J is the total number of models. Since the multimodel average of the IMT estimates $h(t_0)$ is a close approximation to the final multimodel trend estimate (MMT) derived in the third step of the TSAM approach, the baseline adjustment may be viewed simply as forcing the anomaly time series to go roughly through the final MMT estimate at the reference date $t_0$.

[14] The time series (4) contains all the information of (3) plus the multimodel average $h(t_0)$, which can be compared with observations. We will used the baseline $t_0 = 1980$ since a number of the CCMVal-1 models do not have data prior to this date. Following (4), the 1980 baseline-adjusted time series, $y'_{jk}$ for the CCMVal-1 March 60°N–90°N and October 60°S–90°S total column ozone are displayed in Figures 2c and 2d, respectively. The corresponding 1980 baseline-adjusted nonparametric IMT estimates $h'_j(t)$ are presented in Figures 2e and 2f. Following (1) and (4) the 1980 baseline-adjusted nonparametric smooth trend in our model is:

$$h'_j(t) = h_j(t) - h_j(t_0) + h(t_0) \tag{6}$$

with

$$y'_{jk}(t) = h'_j(t) + \epsilon_{jk}(t). \tag{7}$$

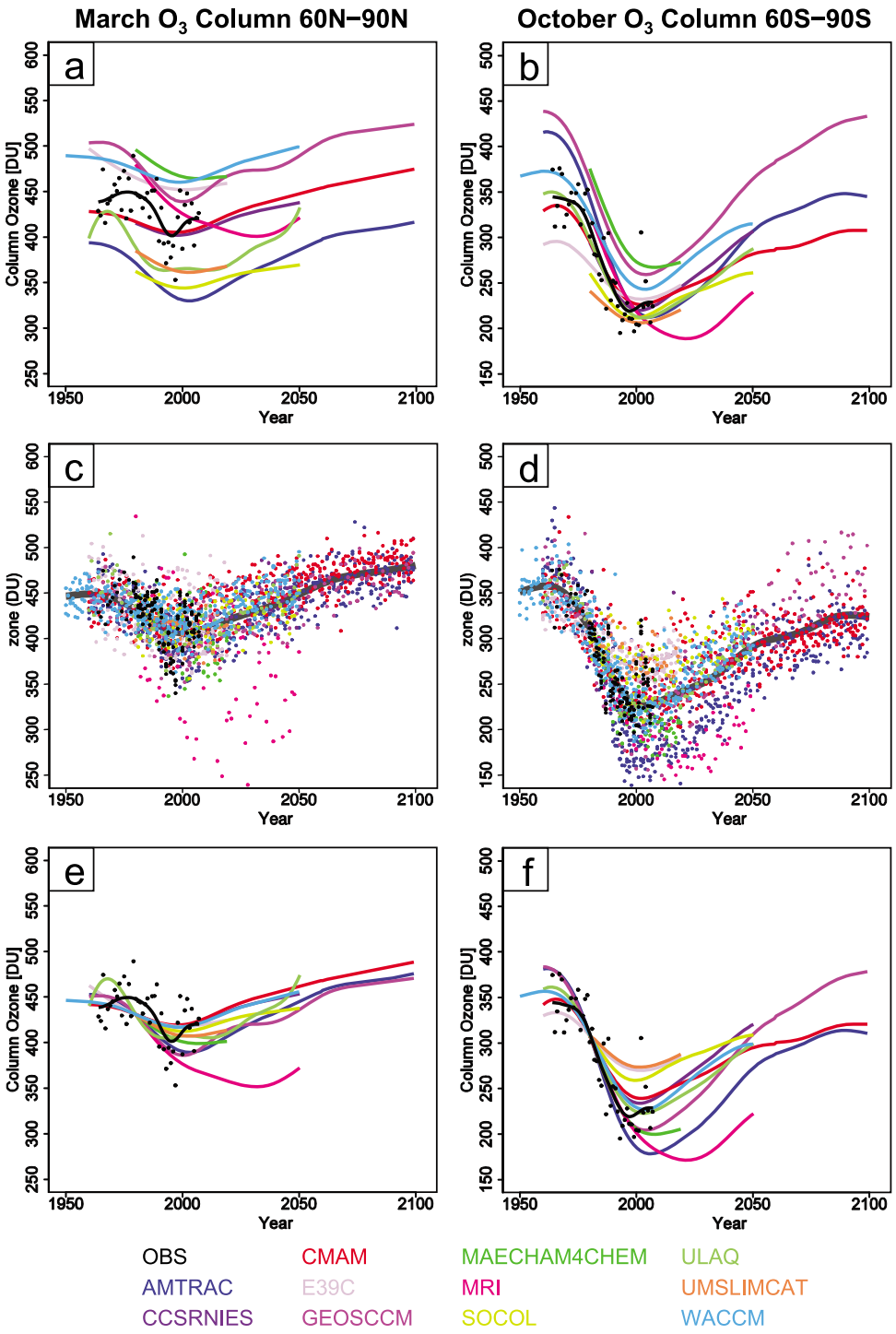


**Figure 2.** (a and b) Initial estimate of the individual model trends $h_j(t)$ for the raw time series displayed in Figure 1. This represents the first step in the TSAM approach. (c and d) The 1980 baseline-adjusted time series data $y'_{jk}$ following from (7) with $t_0 = 1980$. (e and f) The 1980 baseline-adjusted trend estimate $h'_j(t)$. This represents the second step in the TSAM approach. The thick gray line in Figures 2c and 2d represents the trend estimate $g'(t)$ for the simpler nonparametric additive model (8). For reference, following *Eyring et al.* [2007], smooth fits to the observations in these plots have been created by 30 iterations of a 1:2:1 filter (black lines).

Before moving on to the third step in the TSAM, we may ask if the statistical model in (7) is well specified. In other words, are all its model assumptions satisfied in modeling the data, for example, that the noise term $\epsilon_{jk}(t)$ is independent from year to year, is normally distributed, and is drawn from the same underlying distribution with zero mean and similar

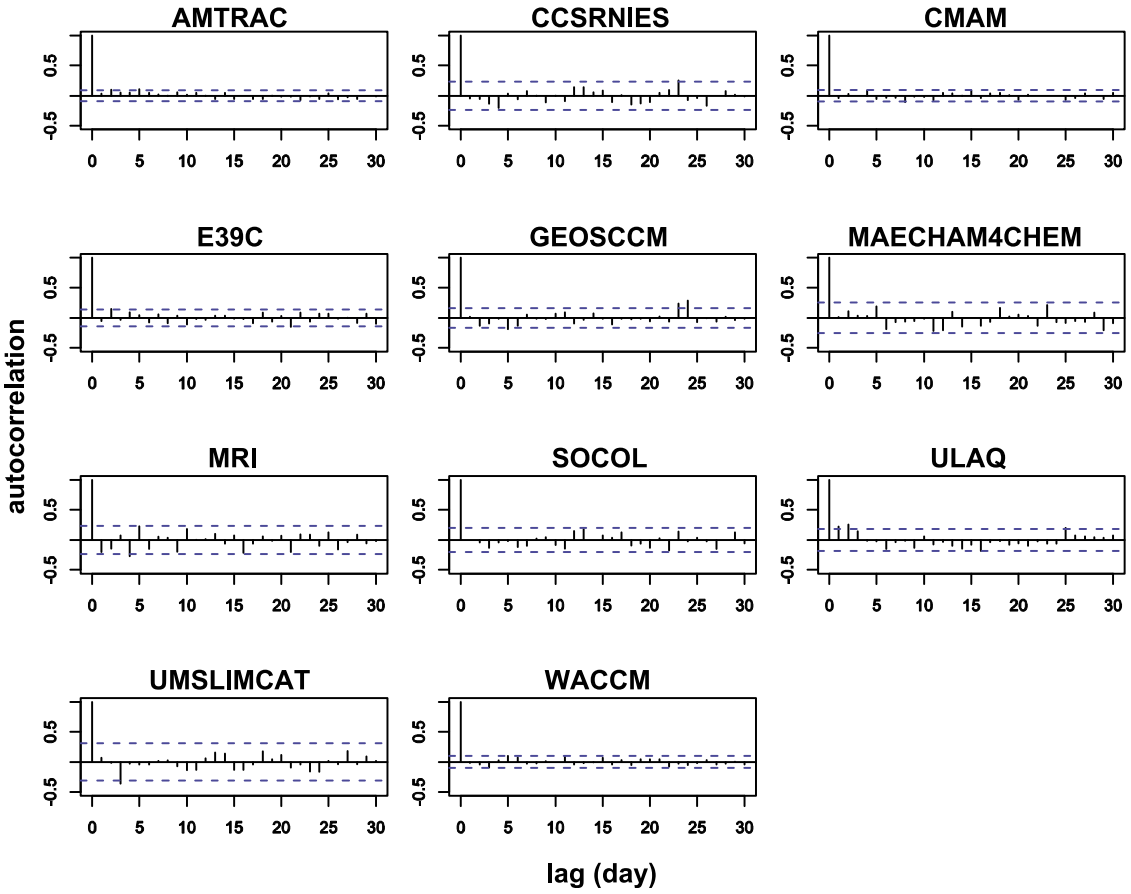


**Figure 3.** Individual model autocorrelation functions for the residuals $\epsilon_{jk}(t)$ for CCMVal-1 October total column ozone in the latitude band 60°S–90°S. This noise corresponds to the nonparametric model (7) with 1980 baseline trend estimates $h'_j(t)$ displayed in Figure 2f. The blue dashed lines represent 95% confidence limits for the sample autocorrelation function. This suggests that the assumption of year-to-year independence is a good one for the (7) model.

variance. For example, we could have chosen the simpler nonparametric model:

$$y'_{jk}(t) = g'(t) + \hat{\epsilon}_{jk}(t), \tag{8}$$

where one trend estimate is made for all time series data instead of individual trend estimates for each model (7). This implicitly defines a different random noise component $\hat{\epsilon}_{jk}(t)$. The nonparametric trend estimate $g'(t)$ is displayed as the thick grey line in Figures 2c and 2d. If (8) were a reasonable model for the data then, in addition to being an IMT, $g'(t)$ could also serve as the MMT thereby eliminating the need for the third step of the TSAM. Visual inspection of the smooth estimate $g'(t)$ to the 1980 baseline-adjusted time series $y'_{jk}$ in Figures 2c and 2d would suggest a reasonable fit. However, because we have built the approach on a probabilistic model, the specification of the $g'(t)$ and $h'_j(t)$ fits may be tested.

[15] The year-to-year independence of the model noise term may be tested by calculating its autocorrelation function. In Figure 3 the autocorrelation function for the noise term $\grave{\epsilon}_{jk}(t)$ is displayed for each model for the nonparametric fit (7) to the CCMVal-1 October 60°S–90°S column ozone. The dashed blue lines in Figure 3 represent 95% confidence limits. Lines that extend beyond these limits are considered to be sample correlations that are significantly different from zero. Inspection of all the models reveals that the assumption of year-to-year independence is a good one for the model (7). The fits are designed to give "smooth" estimates of the long-term trends; in other words, they capture long-term variations in the trend but do not wiggle up and down annually. Interannual wiggles are suppressed by a roughness penalty when fitting the splines. Since the residuals of (7) are not serially correlated, there is no need to represent the residuals using more complex time series models such as ARMA models or fractional Brownian noise. This would not be the case if one had used simple parametric trends that are unable to follow longer term up and down decadal variations. Unlike model (7), the simpler nonparametric model (8), which has the same trend for all climate models, gives serially correlated residuals (Figure 4) and so is not well specified (i.e., its assumptions are not satisfied when the model is fit to the data).

[16] Model assumptions related to the noise term may be further investigated by "notched box-and-whisker" plots. These are displayed for $\hat{\epsilon}_{jk}(t)$ and $\grave{\epsilon}_{jk}(t)$, respectively, in Figures 5a and 5b again for the CCMVal-1 October

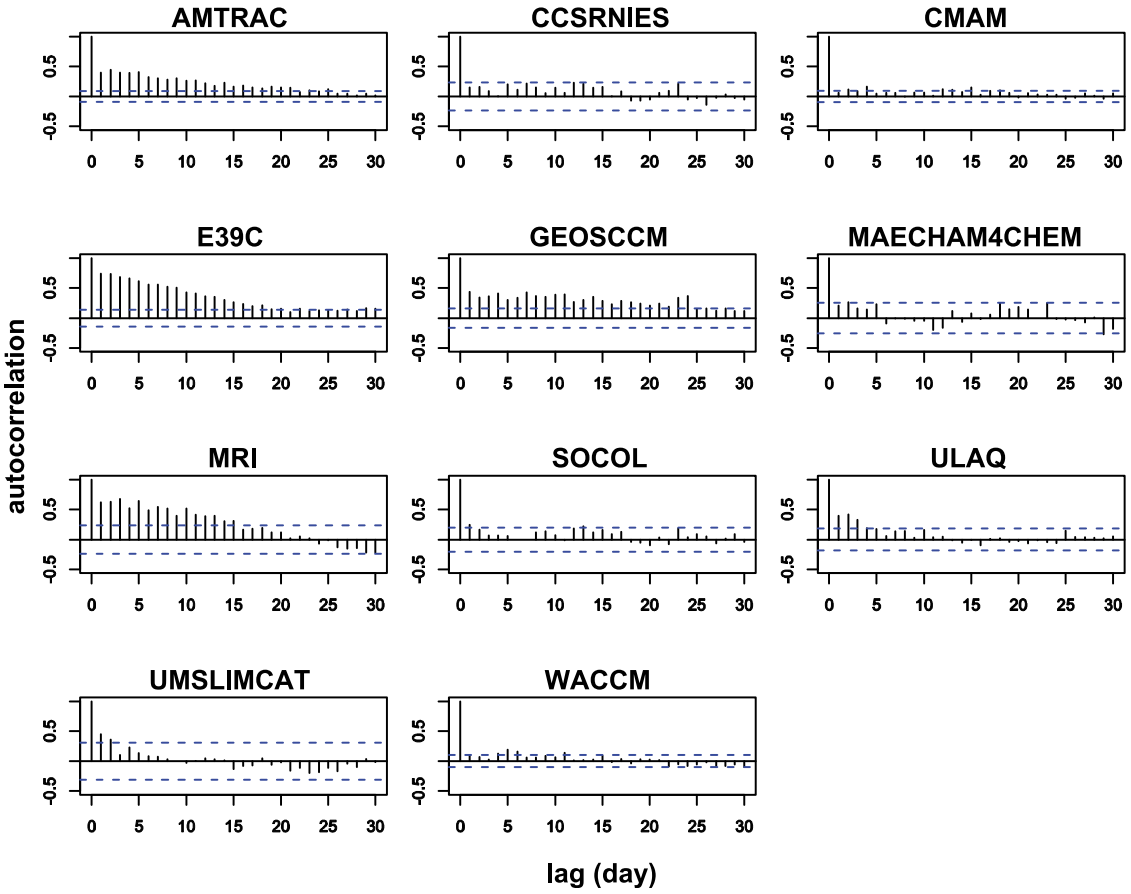


**Figure 4.** Individual model autocorrelation functions for the noise term $\hat{\epsilon}_{jk}(t)$ for CCMVal-1 October total column ozone in the latitude band 60°S–90°S. This noise corresponds to the simpler nonparametric model (8) with a 1980 baseline trend estimate $g'(t)$ displayed in Figure 2d. The lines extending past the blue dashed lines for several models indicates that the assumption of year-to-year independence is not well satisfied for the (8) model.

60°S–90°S column ozone (see caption for details). From Figure 5b, we can see that the noise term $\hat{\epsilon}_{jk}(t)$ has a similar location and scale for each model, validating the model assumption that the residuals were drawn from the same distribution with zero mean and roughly the same variance. Again, the same cannot be said for the $\hat{\epsilon}_{jk}(t)$ residuals (Figure 5a) suggesting that $g'(t)$ in (8) is not a well-specified model of the trend.

[17] We conclude, therefore, that (7) represents one of the simplest nonparametric additive models that is satisfied by the ozone indices considered in the two examples.

### 2.3. Multimodel Trend Estimates

[18] The final step of the TSAM approach involves combining the IMT estimates $h'_j(t)$ to arrive at an MMT estimate:

$$h'(t) = \sum_j w_j(t) h'_j(t), \tag{9}$$

where the weights $w_j(t)$ have the properties

$$w_j(t) \geq 0 \text{ and } \sum_j w_j(t) = 1. \tag{10}$$

If the weights are assumed to be nonrandom, and the errors in the individual trends are assumed to be independent, then the squared standard error of the weighted sum is given by:

$$s_h^2(t) = \sum_j w_j^2(t) s_j^2(t), \tag{11}$$

where $s_j(t)$ is the standard error of the trend estimate $h'_j(t)$, which can be calculated using standard expressions from linear regression [*Woods*, 2006]. The standard error (11) can then be used to estimate the confidence and prediction intervals, respectively, as:

$$\left[h'(t) - 1.96 s_h(t), h'(t) + 1.96 s_h(t)\right] \tag{12}$$

and

$$[h'(t) - 1.96\sqrt{s_h^2(t) + s_\epsilon^2}, h'(t) + 1.96\sqrt{s_h^2(t) + s_\epsilon^2}]. \tag{13}$$

[19] The 95% confidence interval in the trend gives the uncertainty in the trend estimate, where $s_\epsilon$ is the standard deviation of the noise term. In other words, there is 95%

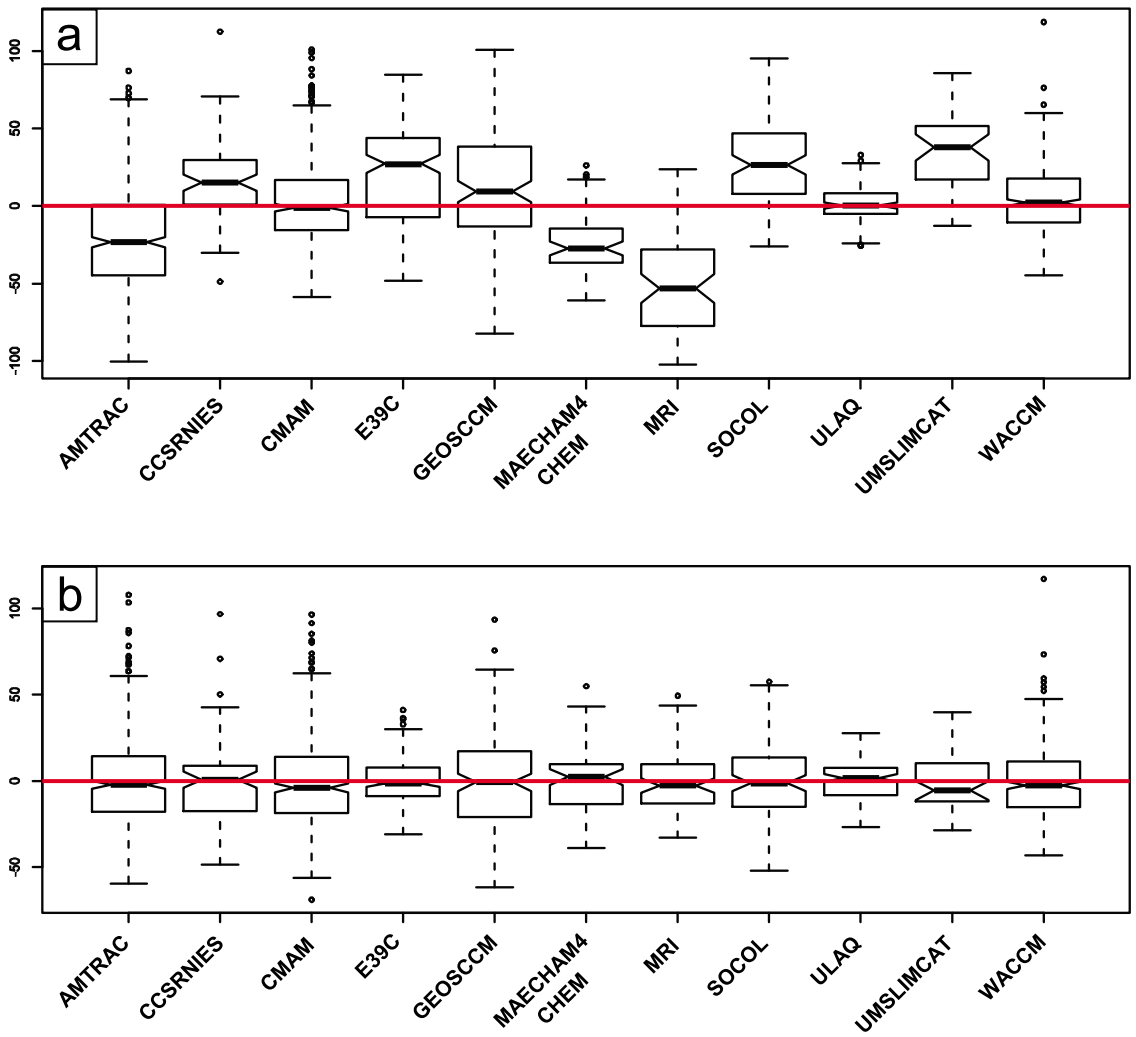


**Figure 5.** Individual model notched box-and-whisker plots (a) for the noise term $\hat{\epsilon}_{jk}$ $(t)$ corresponding to the simpler nonparametric additive model (8) and (b) for the noise term $\epsilon_{jk}(t)$ corresponding to the nonparametric additive model (7). These apply to the CCMVal-1 October total column ozone in the latitude band 60°S–90°S. In these plots the central black line represents the median, the extent of the notches away from the median line indicates the 95% confidence interval of the median, the top and bottom of the boxes represent the upper and lower quartiles, respectively, and the top and bottom whiskers extend out to 1.5 times the distance from the first to third quartiles. For the noise term $\epsilon_{jk}(t)$ in Figure 5b, the medians of all models fall within the notches and are close to zero. Also, the similar height of the boxes indicates that all models have a similar amount of variance away from the estimated trend $h'_j(t)$. For the noise term $\hat{\epsilon}_{jk}(t)$, the means are significantly different and the intermodel variance is larger suggesting that (8) is not a suitable model for these data.

chance that this interval will overlap the expected trend predicted by the statistical model. The interval is pointwise (rather than simultaneous) in that it represents the uncertainty in the trend at each year rather than being an interval for all probable trend curves over the whole period. The 95% prediction intervals give an idea of how much uncertainty there might be in a predicted index value for an individual year. In other words, there is 95% chance that a particular index value on a specific year will lie in this interval. This interval is the combination of uncertainty in the trend estimate and the uncertainty due to natural interannual variability about the trend.

[20] The specific choice of weights in (9) remains open. In general, we decide to base the construction of the weights on a statistical probability model with testable assumptions. Here we have chosen a "random effects" model to determine the weights. This model assumes that the trends for individual models $h'_j(t)$ are random samples from the "true trend" $\tilde{h}_j(t)$:

$$h'_j(t) = \tilde{h}'_j(t) + \eta(t) \tag{14}$$

where

$$\eta(t) \sim N\left(0, \lambda^2\right). \tag{15}$$

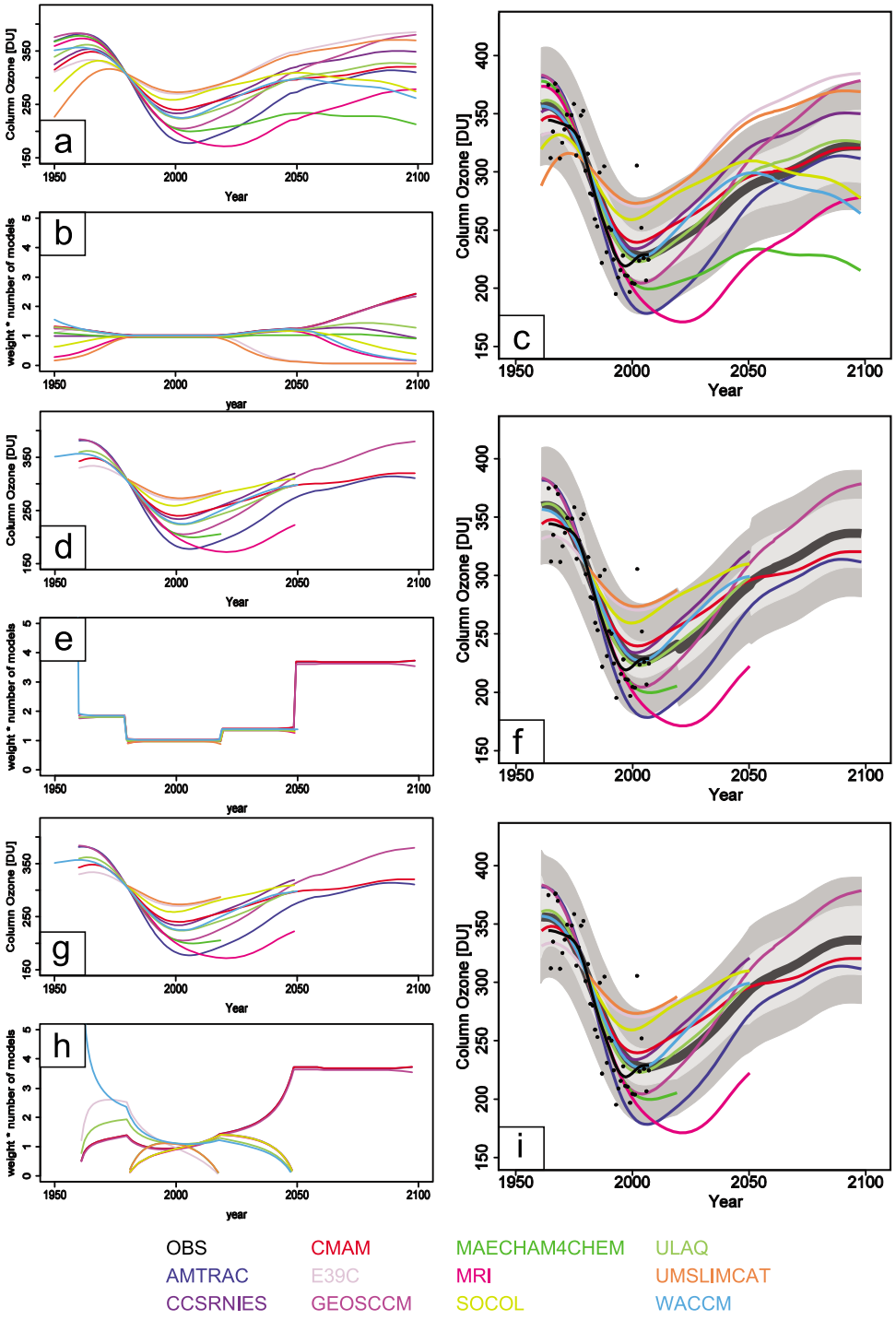


**Figure 6.** For time series of CCMVal-1 October total column ozone in the latitude band 60°S–90°S are presented the (a, d, and g) individual model fits, (b, e, and h) weights, and (c, f, and i; thick gray line) trend (MMT) estimate for three approaches to determining the weights. Results from the "random effects" model (17) are shown in Figures 6a–6c. One problem with this approach is that models can contribute to the final MMT estimate at times when no data exists of that model (i.e., in regions where $h'_j(t)$ represents an extrapolation). The introduction of prior weights (20) can help mitigate this problem. Results from the use of a simple on/off set of prior weights (having a value of one where there are model data and zero where there is none) are presented in Figures 6d–6f. One artifact of this approach is that it causes discontinuities in the final MMT estimate. Finally, results from set of prior weights used for the present study, which employ a smoother quadratic taper from a value of 1 where time series data exist to a value of 0 where it is absent, is displayed in Figures 6g–6i.

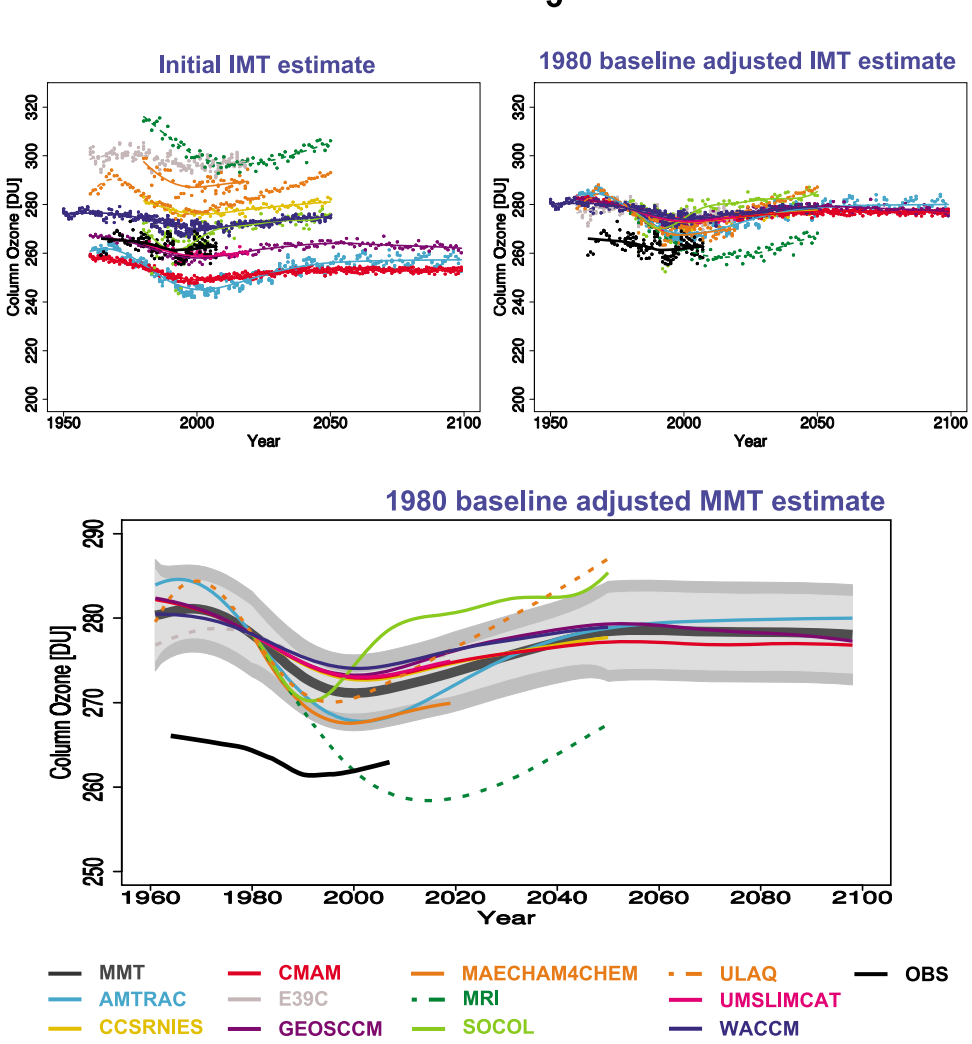


**Figure 7.** Raw time series data of annually averaged column $O_3$ (25°S–25°N) and initial individual model trend (IMT) estimates (top left), and 1980 baseline-adjusted time series data and IMT estimates (top right) for the TSAM analysis of CCMVal-1 data. Observational data (black symbols) and lowess fit with smoother span $f = 0.4$ [*Cleveland and Devlin*, 1988] to the observations appear as black lines in all panels. (bottom) The 1980 baseline-adjusted multimodel trend (MMT) estimate is displayed (heavy dark gray line) with 95% confidence and 95% prediction intervals appearing as light and dark gray-shaded regions about the trend. The 1980 baseline-adjusted IMT estimates are also plotted.

The quantity $\lambda^2$ is included to account for additional variance between model trends that cannot be accounted for merely by sampling the uncertainty $s_j^2$. Using this random effects model, (11) then generalizes to:

$$s_h^2(t) = \sum_j w_j^2(t)\left(\lambda^2 + s_j^2(t)\right), \tag{16}$$

which is used here to calculate intervals. Assuming this model is valid, a least squares estimate of $w_j(t)$ may be obtained from (9) employing the weights:

$$w_j(t) = \frac{w(t)}{\lambda^2 + s_j^2(t)} \tag{17}$$

where

$$w^{-1}(t) = \sum_j \left(\lambda^2 + s_j^2(t)\right)^{-1}. \tag{18}$$

Specification of the weights $w_j(t)$ from (17) requires an estimate of the parameter $\lambda^2$. For this we have used the following iterative approach: An initial estimate of the true trend is obtained by calculating $h'_{\lambda=0}(t)$. Then an iterative Newton-Raphson algorithm is employed to determine the $\lambda$ that gives scaled residuals that have unit variance as is expected from (14):

$$\mathrm{var}\left(\frac{h'(t) - h'_{\lambda=0}(t)}{\sqrt{\lambda^2 + s_j^2(t)}}\right) = 1. \tag{19}$$

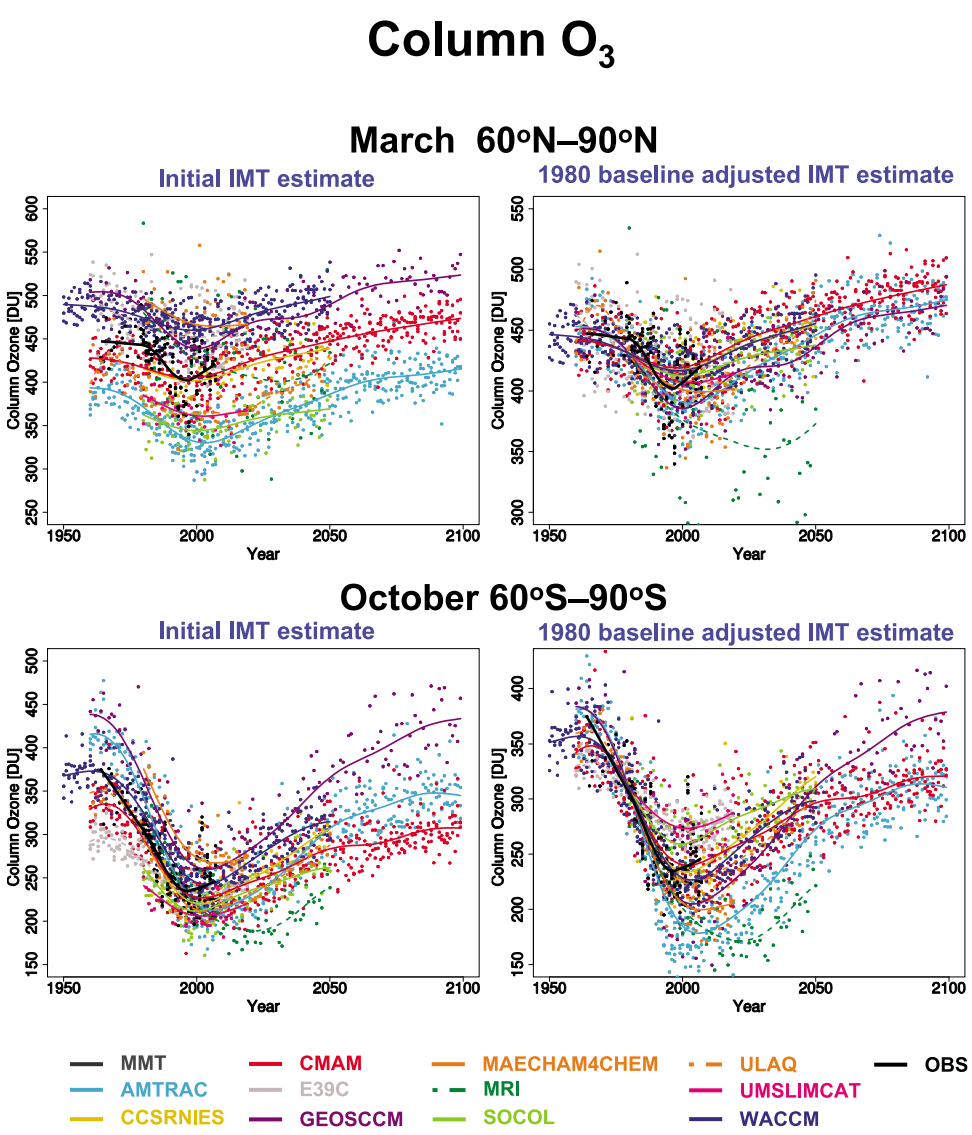


**Figure 8.** Raw time series data of (left) annually averaged column $O_3$ and initial individual model trend (IMT) estimates and (right) 1980 baseline-adjusted time series data and IMT estimates for the TSAM analysis of column ozone in the latitude bands (top) 60°N–90°N and (bottom) 60°S–90°S. Observational data as given as in Figure 7.

Employing this model for the weights produces the MMT estimate $h'(t)$ for the 1980 baseline CCMVal-1 October 60°S–90°S column ozone displayed in Figure 6c. The associated individual model trend estimates $h'_j(t)$ and weights $w_j(t)$ are displayed in Figures 6a and 6b, respectively. In Figure 6, the weights are scaled by the number of models so that a scaled weight of 1 implies a proportional contribution of that model to the MMT estimate.

[21] While this formulation of weights provides a smooth final trend estimate $h'(t)$, for this example it highlights a potential problem: the individual model weights $w_j(t)$ are very insensitive to the absence of data in the original time series. For example, the time series for the MAECHAM4-CHEM model (green) extends only over the period 1980–2019 (see Figure 2). Its scaled weight, however, has a value of roughly 1 over the entire period 1960–2100 suggesting significant contributions of its trend estimate $h'_j(t)$ at times when there are no model data. The original idea behind this model for the weights was that the natural increase in standard errors $s_j^2(t)$ in the region where $h'_j(t)$ is extrapolated beyond the model data would cause the weights to decrease naturally toward zero. While Figure 6b indicates that there is some tendency for the weights to display this behavior, many models retain weight values close to unity out to 2100 where they have provided no data.

[22] To correct this unphysical behavior, we introduce the concept of prior weights $w_j^p(t)$ into the formulation such that the final weights now have the form:

$$w'_j(t) = \frac{w_j^p(t) w_j(t)}{\sum_{j'} w_{j'}^p w_{j'(t)}} \tag{20}$$

(with $w'_j(t)$ implicitly replacing $w_j(t)$ in expressions (11) and (17)). An example set of prior weights would be the "on/off"

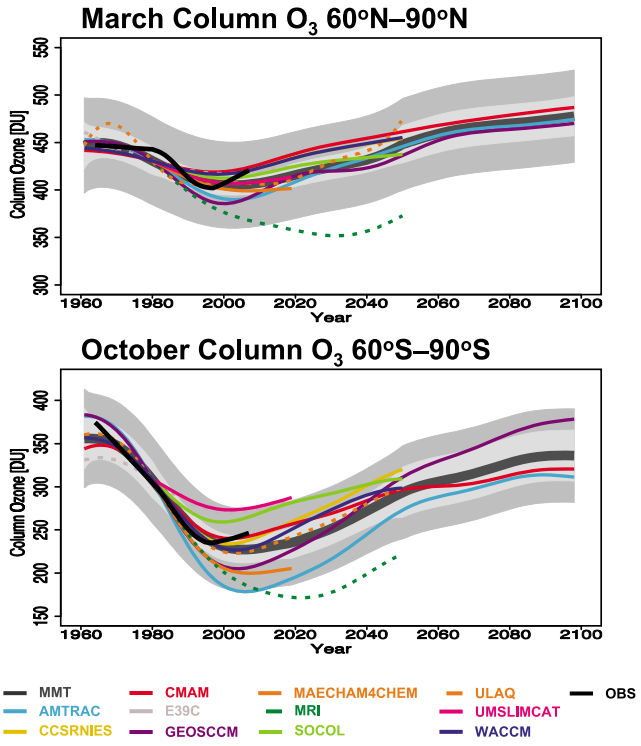


**Figure 9.** 1980 baseline-adjusted multimodel trend (MMT) estimates of annually averaged column $O_3$ (heavy dark gray line) in the latitude bands (top) 60°N–90°N and (bottom) 60°S–90°S with 95% confidence and 95% prediction intervals appearing as light and dark gray-shaded regions about the trend. The 1980 baseline-adjusted IMT estimates and unadjusted lowess fit [*Cleveland and Devlin*, 1988] to the observations are additionally plotted.

set: $w_j^p(t) = 1$ at times $t$ when raw time series data exist for model $j$ and $w_j^p(t) = 0$ otherwise. This prescription is illustrated in Figures 6d–6f. It corrects the unphysical behavior identified when $w_j(t)$ of (17) is used alone. However, this on/off prescription is still problematic in that it causes discontinuities in the MMT estimate Figure 6f. The set of prior weights used for the present chapter employ a smoother quadratic taper from a value of 1 where time series data exists to a value of 0 where it is absent:

$$w_j^p(t) = \begin{cases} 1 - z^2 & \text{if } 0 \le z^2 \le 1 \\ 0 & \text{otherwise} \end{cases}, \tag{21}$$

where

$$z = -1 + 2\left(t - t_{j,\min}\right)/\left(t_{j,\max} - t_{j,\min}\right), \tag{22}$$

and where $[t_{j,\min}, t_{j,\max}]$ defines the period within which data exist for model $j$. This scheme is illustrated in Figures 6g–6i.

[23] Finally, we note that the formulation of prior weights (20) allows a natural entry point for the specification of prior, time-independent, model weights based on performance metrics. Such metric based weights would take on values in the range [0,1] and simply multiply $w_j^p(t)$ in the numerator and denominator of expression (20).

## 3. TSAM Analysis of Column Ozone and Its Recovery

[24] In this section the TSAM approach is illustrated by its application to CCMVal-1 total column ozone in all latitude bands. Individual and multimodel trend estimates derived from this analysis are used to make quantitative estimates of ozone recovery.

[25] In Figure 7 (top left) we present the raw time series and the initial TSAM individual model trend (IMT) estimates for the annual total column ozone in the latitude band 25°S–25°N for 11 CCMVal-1 models. These initial IMT estimates employ the nonparametric additive model discussed in section 2 and were verified by an analysis of the residuals (e.g., see section 2.3). Observations of total ozone from four data sets are also presented in Figure 7 (black lines and symbols). These include ground-based measurements (updated from *Fioletov et al.* [2002]), merged satellite data [*Stolarski and Frith*, 2006], the National Institute of Water and Atmospheric Research (NIWA) combined total column ozone database [*Bodeker et al.*, 2005], and from Solar Backscatter Ultraviolet (SBUV, SBUV/2) retrievals (updated from *Miller et al.* [2002]).

[26] The raw time series display a wide range of background total ozone values over the entire REF-A2 period, which extend significantly above and below the observed values in this region. As described in section 2.2, relative to a selected reference date, baseline-adjusted time series and IMT estimates are computed in the second step of the TSAM approach to facilitate a closer comparison of the predicted evolution of ozone indices between models. Following the analysis performed in chapter 6 of *WMO* [2007] and *Eyring et al.* [2007], anomaly time series are created for each model about a baseline value prior to significant ozone loss. Here the baseline value is taken to be the initial IMT estimate at a selected reference date for each model (e.g., 1980). The baseline adjusted time series are then formed by adding a constant so that each anomaly time series goes through the multimodel average of the IMT estimates at the reference date. Since the multimodel average of the IMT estimates is a close approximation to the final multimodel trend estimate (MMT) derived in the third step of the TSAM approach, the baseline adjustment may be viewed simply as forcing the anomaly time series to go roughly through the final MMT estimate at the reference date.

[27] The baseline-adjusted IMT estimates employing a reference date of 1980 are presented in Figure 7 (top right). Comparing the top left and top right of Figure 7, it can be seen that the TSAM approach has been very effective at providing a common reference for the total ozone time series allowing a clearer comparison of the predicted evolution between models. In Figure 7 (bottom) the multimodel trend (MMT) estimate (thick gray line) computed in the final step of the TSAM approach for the 25°S–25°N total column ozone is presented. The 95% confidence and 95% prediction intervals for the MMT estimate are also displayed as the light and dark gray-shaded intervals and the IMT estimates are superposed on top of the MMT estimate.

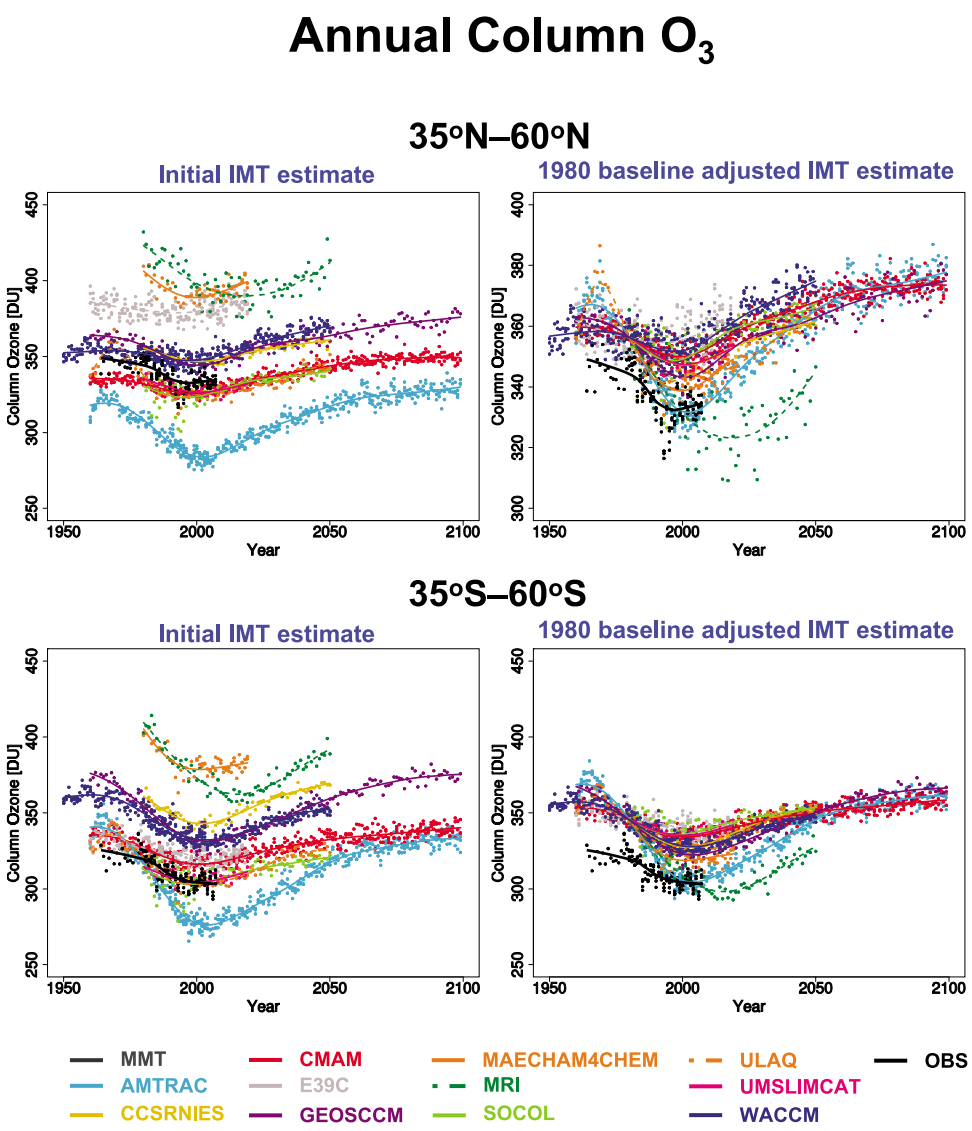


**Figure 10.** Same as Figure 8 but for the annual mean column ozone in the latitude bands (top) 35°N–60°N and (bottom) 35°S–60°S.

[28] The TSAM analysis of springtime total column ozone over polar latitudes (60°N–90°N in March and 60°S–90°S in October) is presented in Figure 8. Unlike the tropical latitudes, the baseline adjustment of the column ozone has only a modest impact on the intermodel spread in the Arctic and essentially no impact in the Antarctic. This implies that tropical differences come primarily from intermodel systematic offset biases while polar differences come from intermodel differences in sensitivity to ozone depleting substances. The MMT estimates of the polar ozone are displayed in Figure 9. Due to the greater intermodel spread of the baseline adjusted time series at the poles relative to the tropics, there occurs larger 95% confidence and prediction intervals about the MMT estimates in the polar latitudes. It is important to note that, while the TSAM analysis has been effective at combining IMT estimates into the final MMT estimate, an artifact in the form of a change in slope seems to remain near 2050 in Figure 9. This is primarily due to the fact that a large number of the model time series end at this point and the MMT estimate beyond this time relies on only three models' IMT estimates.

[29] The TSAM analysis of annual column ozone in midlatitudes (35°N–60°N and 35°S–60°S) is presented in Figure 10. The baseline adjustment in both the Northern Hemisphere and Southern Hemisphere is found to reduce the intermodel spread of the IMT estimates. The MMT estimates of the midlatitude ozone are displayed in Figure 11. Due to the reduced intermodel spread of the baseline adjusted time series at midlatitudes relative to the poles, there occurs tighter 95% confidence and prediction intervals about the MMT estimates in midlatitudes.

[30] The IMT and MMT estimates for total ozone may be used to quantify individual model, and multimodel estimates of ozone recovery back to values associated with a specified

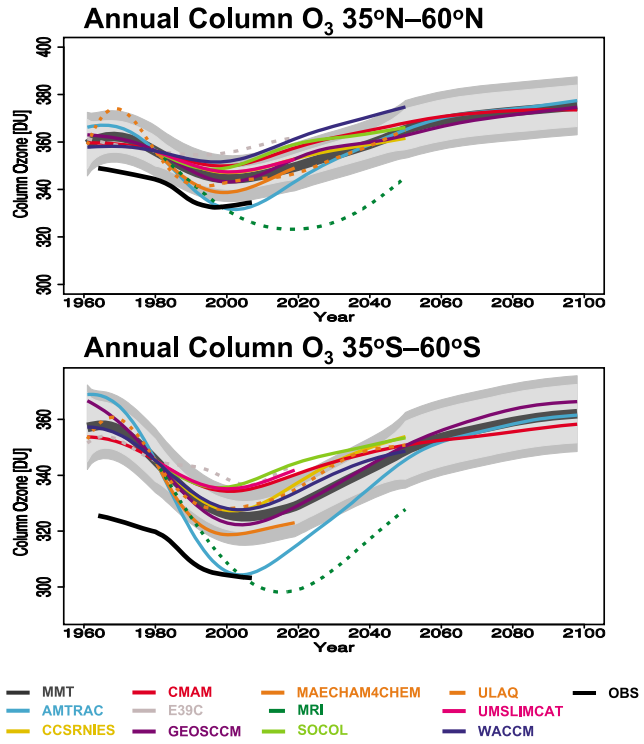


**Figure 11.** Same as Figure 9 but for the annual mean column ozone in the latitude bands (top) 35°N–60°N and (bottom) 35°S–60°S.

reference date. Because the IMT and MMT estimates are smooth curves by construction, the value of ozone for any reference date prior to maximum ozone depletion may be mapped onto a future date based on the return of ozone to the reference date value. The TSAM approach, therefore, allows the definition of return dates for a continuous set of reference dates. Here a reference date of 1980 will be used for the REF-A2 time series to be consistent with previous analyses of ozone return dates for the CCMVal-1 data set [*WMO*, 2007, chapter 6; *Eyring et al.*, 2007].

[31] In Figure 12, for the CCMVal-1 data set, we present summary diagnostics of 1980 return dates for column ozone for the five sets of latitude bands considered in Figures 7–11. In Figure 12, for each latitude band, the MMT estimates of return dates are indicated by large black triangles. Error bars on these estimates are associated with the TSAM 95% confidence intervals. Figure 12 provides a concise summary of the ozone evolution presented in Figures 7–11. It allows return dates of individual models to be compared to each other, and to the return date of the multimodel mean. It also clearly reveals the latitudinal structure of recovery.

[32] Initial inspection of Figure 12 reveals that return dates for total ozone are not symmetric in latitude and in the tropics realized by only a few models, with the MMT estimate of return date having a very large uncertainty. The asymmetric structure of midlatitude and polar ozone recovery is an indication that, in addition to halogen loading, ozone is affected by dynamical and radiative changes brought about by increased greenhouse gas forcing (e.g., *Austin et al.* [2010]) and these have been consistently reproduced in the MMT estimates of CCMVal-1. In the absence of the TSAM analysis this result was difficult to identify.

## 4. Summary and Discussion

[33] In this study we have developed a nonparametric additive model trends analysis that is suitable for producing smooth individual model trend (IMT) and multimodel trend (MMT) estimates of time series of unequal length. We have referred to this approach as the Time Series Additive Model, or "TSAM," approach. The association of the TSAM approach with a probabilistic model allows the trend estimates to be used to make formal inference (e.g., calculation of confidence and prediction intervals). Advantages of the TSAM include: the production of smooth trend estimates out to the ends of the time series, the ability to model explicitly interannual variability about the trend estimate, and the ability to make rigorous probability statements.

[34] The TSAM approach was motivated by the nature of the CCMVal-1 data set, which formed the basis of the 2006 Ozone assessment [*WMO*, 2007]. In general, individual model time series of future ozone in the REF-A2 experiment of CCMVal-1 spanned only portions of the requested period (1980–2100) making it difficult to evaluate multimodel ensemble behavior. The TSAM approach was illustrated by its application to REF-A2 ozone time series providing smooth IMT and MMT estimates that extended over the entire REF-A2 period. This allowed, for the first time, quantitative estimates of ozone return dates for individual models and the multimodel mean of the CCMVal-1 data set.

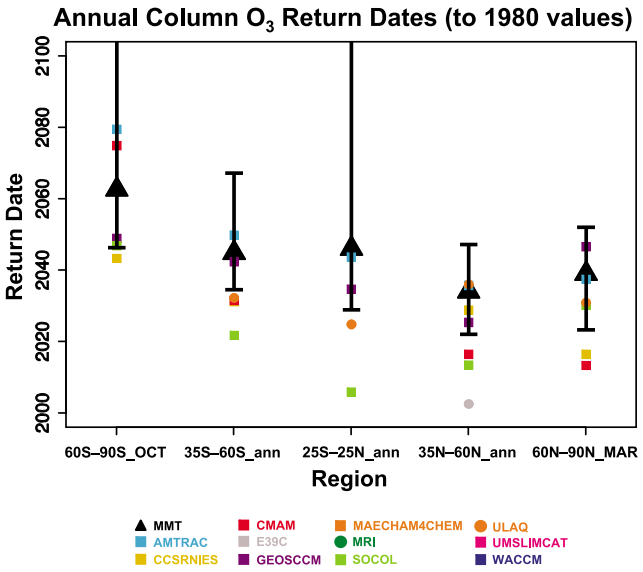


**Figure 12.** Date of return to 1980 values for the annual average (tropical and midlatitude) and spring (polar) total ozone column derived from the IMT (colored symbols) and MMT (large black triangles) estimates for the CCMVal-1 data set. The error bars on the MMT estimate of recovery date are derived from the 95% confidence interval of the MMT estimates to the 1980 baseline-adjusted time series data.

[35] It is anticipated that as the number of models increases, assuming all span the same period, a simple multimodel ensemble mean and the MMT estimate of the TSAM approach will converge toward each other. Smooth estimates of individual model trends, on the other hand, would require ensembles of nontrivial size for each model. In the absence of large ensembles for each model, the IMT estimate of the TSAM would seem to provide a superior estimate of smooth individual model trends.

[36] The TSAM approach can easily be extended to time series with shorter sampling times. For monthly (or even daily) time series one is likely to have to model serial correlation in the residuals. A good example of how to use autoregressive models to do this for daily air temperatures in Cairo is presented on page 322 of *Woods* [2006].

[37] Finally, since one of the main advantages of the TSAM is the use of the probabilistic model, the approach provides both the mean model trend and an estimate of its uncertainty. Although this uncertainty does not take account of unknown physical processes, ozone in the atmosphere is generally well understood. Therefore, it is anticipated that the statistical uncertainty in the multimodel trend is representative of the true uncertainty, and that its knowledge would be of benefit in the formulation of policy on the ozone layer.

[38] **Acknowledgments.** The authors would like to thank T.G. Shepherd for helping to initiate this collaboration, and DBS would like to thank Simon Woods for a useful discussion about standard errors in GAM models. J.F.S. has received support in part from the Canadian Foundation for Climate and Atmospheric Sciences through the C-SPARC network, and J.A.'s research was administered by the University Corporation for Atmospheric Research at the NOAA Geophysical Fluid Dynamics Laboratory.

## References

Andersen, S. B., et al. (2006), Comparison of recent modeled and observed trends in total column ozone, *J. Geophys. Res.*, *111*, D02303, doi:10.1029/2005JD006091.

Austin, J., et al. (2010), Decline and recovery of total column ozone using a multimodel time series analysis, *J. Geophys. Res.*, *115*, D00M10, doi:10.1029/2010JD013857.

Bodeker, G. E., H. Shiona, and H. Eskes (2005), Indicators of Antarctic ozone depletion, *Atmos. Chem. Phys.*, *5*, 2603–2615.

Cleveland, W. S., and S. J. Devlin (1988), Locally weighted regression: An approach to regression analysis by local fitting, *J. Am. Stat. Assoc.*, *83*, 596–610.

Eyring, V., et al. (2006), Assessment of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past, *J. Geophys. Res.*, *111*, D22308, doi:10.1029/2006JD007327.

Eyring, V., et al. (2007), Multimodel projections of ozone recovery in the 21st century, *J. Geophys. Res.*, *112*, D16303, doi:10.1029/2006JD008332.

Eyring, V., M. P. Chipperfield, M. A. Giorgetta, D. E. Kinnison, E. Manzini, K. Matthes, P. A. Newman, S. Pawson, T. G. Shepherd, and D. W. Waugh (2008), Overview of the new CCMVal reference and sensitivity simulations in support of the upcoming ozone and climate assessments and the planned SPARC CCMVal, *SPARC Newsl.*, *30*, 20–26.

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

Hofmann, D. J., S. Oltmans, J. Harris, B. Johnson, and J. Lathrop (1997), Ten years of ozonesonde measurements at the south pole: Implications for the recovery of springtime Antarctic ozone, *J. Geophys. Res.*, *102*(D7), 8931–8943, doi:10.1029/96JD03749.

Miller, A. J., et al. (2002), A cohesive total ozone data set from SBUV(/2) satellite system, *J. Geophys. Res.*, *107*(D23), 4701, doi:10.1029/2001JD000853.

Newchurch, M. J., E.-S. Yang, D. M. Cunnold, G. C. Reinsel, J. M. Zawodny, and J. M. Russell (2003), Evidence for slowdown in stratospheric ozone loss: First stage of ozone recovery, *J. Geophys. Res.*, *108*(D16), 4507, doi:10.1029/2003JD003471.

R Development Core Team (2008), R: A language and environment for statistical computing, report, R Found. for Stat. Comput., Vienna, Austria. (Available at http://www.R-project.org)

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

Weatherhead, E. C., et al. (2000), Detecting the recovery of total column ozone, *J. Geophys. Res.*, *105*(D17), 22,201–22,210, doi:10.1029/2000JD900063.

Woods, S. N. (2006), *Generalized Additive Models: An Introduction with R*, Chapman and Hall, London.

World Meteorological Organization (WMO) (2003), Scientific Assessment of Ozone Depletion: 2002, *Rep. 47*, Global Ozone Res. and Monit. Proj., Geneva, Switzerland.

World Meteorological Organization (WMO) (2007), Scientific Assessment of Ozone Depletion: 2006, *Rep. 50*, Global Ozone Res. and Monit. Proj., Geneva, Switzerland.

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

J. Austin, University Corporation for Atmospheric Research/Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, NJ 08542-0308, USA.

J. F. Scinocca, Canadian Centre for Climate Modeling and Analysis, Environment Canada, P.O. Box 3065 STN CSC, Victoria, BC V8W 2Y2, Canada. (john.scinocca@ec.gc.ca)

D. B. Stephenson and T. C. Bailey, Mathematics Research Institute, University of Exeter, Exeter EX4 4QF, UK.