The authors wish to thank reviewer #1 for their review.

General Comments:

The fact that the stratosphere is a key driver of tropospheric ozone variability in many locations stands in contrast to the emission centric viewpoint of tropospheric ozone variability that is often prevalent in the literature. As such we feel that this paper strengthens, extends and complements the arguments recently made in recent papers on the global scale (e.g., Hess and Zbinden, 2013; Neu et al., 2014) and the regional scale (e.g., Lin et al. (2012), Ordonez et al. (2007), Tarasick et al. (2005)).

Specific comments (reviewers comments in italics):

1. **Reply:** We agree that the proposed title change would be more appropriate.

   **Change:** We will change the title as suggested by the reviewer: “Ensemble Simulations of the Role of the Stratosphere in the Attribution of Northern Extra-Tropical Tropospheric Ozone Variability”.

2. **Reply:** In the revised version we will change the sentence: “Despite the simplicity of the tropospheric chemistry, the FR WACCM simulations capture the measured N. H. background interannual tropospheric ozone variability in many locations to a surprising extent, suggesting the importance of external forcing in driving interannual ozone variability”.

   **Change:** “Despite the simplicity of its tropospheric chemistry, at a number of N.H. locations the interannual ozone variability in the FR WACCM simulations is significantly correlated with the measured interannual variability, suggesting the importance of external forcing in driving interannual ozone variability”

3. **Reply:** In the revised version we will change the sentence: “The ensemble average tropospheric ozone record can largely be explained as a linear combination of the 30–90 area averaged 150hPa ozone flux and the global methane concentration. “

   **Change:** “The ensemble average tropospheric ozone record can largely be explained as a linear combination of the 30–90° area averaged flux of ozone across the 100 hPa surface and the global methane concentration. “

   Note: We actually calculated the ozone flux across the 100 hPa surface, not the 150 hPa surface. This will be corrected here and throughout the text.

4. **“I think I largely agree with your caveats concerning the use of a simplified tropospheric chemistry scheme. I wonder if your simple scheme has a significantly different tropospheric ozone lifetime compared to a more complex scheme, and whether this may be important for the downwards propagation of ozone anomalies? For example, your simple CH4-NOx scheme presumably doesn’t represent PAN, and thus misses some long-range transport of NOx. I find it hard to gauge how important this might be. Have you compared local ozone lifetimes in this model to your results in Hess and Zbinden (2013), with a more comprehensive tropospheric chemistry scheme? This may be instructive. I am slightly worried that by using a simple tropospheric chemistry you significantly change the lifetime of ozone in the troposphere, and thus either over- (or possibly even under-) emphasize the role of stratospheric ozone relative to in-situ production.”**
**Reply:** The reviewer is correct in their assumption that the simple CH4-NOx scheme does not represent PAN, and thus misses some of the long-range transport of NOx (although NOx itself likely has a sufficiently long lifetime in the upper troposphere that even without the formation of PAN intercontinental transport of NOx is likely to be significant).

We would have liked to compare lifetimes with a more comprehensive model. However, the WACCM simulations we analyzed do not have the required output (i.e., the photochemical loss of ozone). These simulations were originally designed with a stratospheric focus.

**No Change.**

5. **Reply:** The ensemble members were initialized from different initial conditions.

**Change:** We will include this in the revised text.

6. **Just a grumble really: I do find figures etc. in a supplement annoying – I only printed out the main paper and not the supplement, so I have ignored them.**

**Reply:** Yes, I agree this can be annoying. The second reviewer also felt that we should include additional figures in the text.

**Change:** We will include the supplementary figures in the main text.

7. **It is interesting that ozone data over Japan cannot be successfully simulated. Is all the Japanese data unreliable? Similarly for the European data prior to 1990 (or 1998). It seems a bit convenient/sweeping to discount all this data. Could you expand on why this data is considered unreliable?**

**Reply:** One test for reliability might be how well the records of closely spaced stations are correlated. Over Europe there is a rather poor correlation between the measurement stations in the troposphere, even the ones that are quite close together. This raises some question about the accuracy of the measurement record. Hess and Zbinden (2013) pointed this out. Using a similar methodology of comparing the record between closely spaced stations Logan et al. (2012) also raises questions about the accuracy of European ozonesondes prior to 1998. In particular the ozone measurements from the MOZAIC program (available after 1994) and the measurements at the surface European alpine sites (e.g., Jungfraujoch and Zugspitze) are judged to be the most accurate. It is not clear why the earlier ozonesonde measurements are unreliable although Logan et al. (2012) lists a number of possible causes. Thus over Europe the earlier ozonesonde data should be treated with a great deal of care. The alpine site data is likely more reliable, although Logan et al. (2012) points out some possible glitches.

At 150 hPa the Japanese stratospheric ozonesonde measurements show very poor temporal correlation although the tropospheric measurements are better correlated (Hess and Zbinden, 2013). The poor correlation at 150 hPa may or may not be due to poor measurement reliability. We note that the latitudinal span of the Japanese stations is quite large (32°-43°N), possibly contributing to the poor correlation. We note that in both the stratosphere and the troposphere the ozone record for the Japanese stations shows rather poor temporal correlation with the record at many of the other more Northerly locations examined. Examining the EOF pattern, the first EOF has a comparatively small amplitude over the Japanese stations suggesting the processes that control Japanese ozone variability may be rather different from those that control much of the
larger region from 30–90° N. This is likely consistent with the more southerly location of the Japanese stations. At this point we don’t know the cause of the poor correlation between the measurements and the simulation: it is possible the station record is unreliable, or it is possible the model does not accurately capture the interannual variability over Japan.

**Change:** We will change the text to better reflect the fact that only the 150 hPa interannual ozone record over Japan shows little correlation between measurement sites.

8. *The sentence beginning ‘To’ is unclear to me. Do you mean: ‘If the model variability arose purely due to internal model dynamics, we would expect the ozone records from the different ensemble members to be uncorrelated with each other and uncorrelated with the measurements.’?*

**Reply:** Yes. This is what we mean.

**Change:** We will reword the sentence to: “To the extent that the ozone record is driven by internal model dynamics versus interannually varying external forcing we would expect the ozone records from the different ensemble members to be uncorrelated with each other and uncorrelated with the measurements”

9. **Reply:** Agreed.

**Change:** We will write “externally forced” here.

10. *The lower correlations nearer the surface could be related to the shorter ozone lifetime closer to the surface. I think you say as much later it may be worth mentioning here.*

**Reply:** Agreed.

**Change:** In the revised version we will include: “These lower correlations likely reflect the shorter lifetime of ozone near the surface and thus an increased importance of local and regional processes.”

11. *Comments on Figure 7: I wonder if the trend in wstar could (at least partly) reflect an expansion in the latitudinal width of tropical upwelling (and shrinkage in extratropical downwelling)? I.e. is 30–90N all downwelling at 150 hPa, or does it include some times/regions of upwelling? (e.g., during NH summer.)*

Reply: There are indeed times when there is tropical upwelling north of 30° N. During the course of the simulation both the rate of downwelling increases (averaged over those regions where there is downwelling; Figure 1 below) and the area of downwelling increases (Figure 2 below). Other modeling studies have found a narrowing in the upwelling branch of the Brewer-Dobson circulation (Li et al., 2010; Hardiman et al., 2014), consistent with a widening in the downwelling branch.
Figure 1: Zonally averaged w* (m/s) weighted only over those regions where it is downwards from 30-90°N for the four WACCM ensemble simulations. Monthly averaged values are zonally averaged and smoothed over 12 months.

Figure 2: Fraction of area that zonally averaged w* is downwards from 30-90°N for the four WACCM ensemble simulations. Monthly averaged values are zonally averaged and smoothed over 12 months.

**Change:** We will include in the text that the simulation change in w* is due to both strength in w* over the downwelling region and an increase in area of the downwelling region.
12. – is the cubic fit with respect to time or CH4?

Reply: The fit is against methane.

Change: We will clarify this in the following sentence: “We find that a cubic fit of ozone against methane captures some of the subtleties of the long-term record better than a quadratic fit.”

13. ‘The long-term cubic fit to simulated ozone is not linear...’ This seems a bit oxymoronic? If the cubic fit was linear, I’d hope you would have just used a linear fit, not a cubic?

Reply: Good point.

Change: We will delete this sentence.


Change: We will reword this sentence to: “After 1985 the shape of the various measured curves is not consistent with the emission changes”

15. I think you are missing the word ‘flux’, which makes this confusing – i.e. I think you mean 150 hPa ozone flux.

Change: Yes it should be ozone flux. We will correct in the revised version.

16. I wondered what the 5-6 months lag meant physically. Is this the ozone lifetime at 150 hPa? Or the (mean?) transport timescale from 150 hPa to the surface? I would like to understand the significance of this lag length.

Reply: Using a scale height of 7 km, 150 hPa corresponds to approximately 13 km and 300 hPa corresponds to a height of approximately 8.5 km. We assume 300 hPa corresponds to the height of the tropopause. Given an average downward velocity of approximately 3.4 10^{-5} m/s (see figure above) in the lower stratosphere this gives a timescale of: (13-8.5) (km) x 1000 (m/km) /3.4 10^{-5} (m/s) ~ 153 days. Below the tropopause (e.g., ~300 hPa) the mixing is likely to be more rapid. Thus the timescale of 5-6 months is consistent with a lower stratospheric advective timescale. Reviewer #2 also raised additional questions about this timescale and we have replied in some more depth there.

Change: We will include in the text that the lag is consistent with lower stratospheric advective velocities.
17. **Suggest insert partly or largely before ‘ascribed’**

**Change:** Thank you.

18. *The values are 0.84 and 0.73 in Table 3 (i.e. different to those in text).*

**Change:** The values in the text 0.83 and 0.74 should be 0.84 and 0.73 (consistent with the table). I inadvertently switched the second digit and will correct.

19. *‘physically deep’ – do you mean they span from 150 hPa to the surface?*

**Change:** Yes, we will clarify this by qualifying physical deep as spanning from at least 150 hPa to the surface.

20. *By large you mean ~0.4 ppb?*

**Change:** We will change “large” to “comparatively large”

21. *‘...high alpine sites over Europe...where amplitude of first EOF is also large.’ I can’t see high values over the Alps, either at the surface or 500 hPa in Figure 11. Clarify what you mean.*

**Reply:** These measurements are usually made in free tropospheric air. Thus it is appropriate to compare the surface ozone concentrations measured at these mountain sites with simulated ozone concentrations at the altitude of the measurements, above the model surface. While the surface EOF values at the surface are small over Europe, at 500 hPa (Figure 11) the alpine sites show a sizeable EOF response of close to 1 ppbv.

**Change:** This will be clarified in the revised version.

22. *I think these locations can also be usefully described as storm tracks?*

**Reply:** They are certainly associated with the storm tracks, but I would be hesitant to equate the two.

**No change.**

23. *This suggests the external forcing is sometimes important, sometimes not?*

**Reply:** It suggests the response is likely sensitive to the particular kind of forcing. This may have to do with the particular type of “El Nino” event or the conjunction between El Nino and the QBO.

**Change:** We will clarify this in the revised paper.

24. *Figure 3 (and subsequent related figures): Is the model line all stations, all of the*
time, or does it mimic the observations used? I guess it must be all + all, as it goes back beyond the first measurement data. Is this subtlety important?

Reply: We show the monthly averaged simulated ozone at all the stations within a region. Going backwards in time there are fewer measurements made at fewer stations. This makes it increasingly likely that measurements will not represent the regional variability. We do not have the option of sampling at the measurement sampling times as only monthly averaged output is available from the model simulation. Would our correlations improve if we sampled only at the points where ozone is measured geographically? It is possible, but it seems unlikely. (1) The correlation between the measurement sites within each region is generally high in the simulations. It seems unlikely that sampling the model at only those locations where ozone was actually measured during any given month would improve the model’s correlation with the measurements. (2) Since these are free running simulations it seems unlikely the simulations will have much skill in capturing finer scale regional features, e.g., the geographic differences in the interannual variability of ozone within a region. So as to best represent simulated interannual ozone variability as sampled over current sampling network we choose to retain the current sampling methodology.

Change: We will clarify that the simulations are not sampled according to the temporal sampling variability at the individual measurement sites.

24. Technical corrections

Change: Thank you. We will fix these in the revised version.

References:


The authors wish to the reviewer for their review.

General Comments (reviewers comments are in italics)

Primary concerns:
1. Several of the results should be compared/contrasted in more detail with previous studies. In particular, the magnitude of the stratospheric impact on the tropospheric ozone found in this work appears to be much greater than what is found by Neu et al. (2014) in their observational based study.

Reply: This is a good suggestion. Neu et al. (2014) found the correlation between 40-50°N 150 hPa O3 and 30-50°N 510 hPa O3 to be 0.40. As we state in the paper: “The correlation between the 150hPa area averaged ozone and the detrended tropospheric area average at 500hPa reaches 0.80 with a lag of 3months; the correlation between 150hPa and surface ozone reaches 0.75 with a lag of 4months.”

Reasons for these differences are unclear but might include the length of the sample compared, model biases, measurement biases, and the somewhat different quantities compared.

In the conclusions we state: Extrapolating these changes, a 30% increase in the ozone flux by 2100 (Hegglin and Shepherd, 2009) would result in 3% increase in surface ozone and a 6% increase in 500hPa ozone. The analysis of Neu et al. (2014) gives a 40% increase in the ozone flux would change tropospheric in Northern mid-latitudes by 2%. While Neu et al. (2014) does not specify the exact level that this change will occur at it is clear our paper is at least a factor of two higher.

Change: Results in the revised paper will be explicitly compared with Neu et al (2014).

2. Also the ozone flux across the 150 hPa surface (1.1 x 10^4 kg/yr; Figure 7) is significantly less than current estimates of stratosphere-troposphere exchange of ozone (For example, the model study by Hsu and Prather (2009) and the observation/model study by Olsen et al. (2013)). I first assumed these results were per unit area and just missing the units, such as km-2. Even if so, the reported values are still much less than current estimates.

Reply: Thanks for catching this. The figure is correct but with the units kg/sec (the caption says kg/year which is wrong and axis simply gives kg, which is also not correct). The revised version will give Tg/year in the figure.

Note also that the actual calculation used in the paper is the ozone flux across the 100 hPa surface, not the 150 hPa surface as stated in the text. The global flux of ozone across the 100 hPa surface in these simulations is 513 Tg/year. The mean ozone flux from 30-90°N is 352 Tg/year. These estimates are consistent with current estimates.
Change: The revised paper will show the figure in Tg/year and will note the global total. In addition the revised paper will state the vertical ozone flux is calculated at 100 hPa.

3. Much more discussion needs to be made of the lag times used in the correlations. The lag times were selected by minimizing chi squared. However, these considerable lags (e.g. Table 3) do not seem to have any physical justification. The tropospheric lifetime of ozone is thought to be on the order of a month (e.g. Stevenson et al., 2006). If a mass of ozone descends into the troposphere, I would expect most of those ozone molecules to be lost by 5-6 months later. Later in the paper (p20486), the first EOF is associated with regions of known stratosphere-troposphere exchange, particularly deep exchange. The works cited do demonstrate these preferred regions but they also demonstrate that the exchange is relatively rapid. This does not support the justification of 6-9 month lags used for the correlations in Table 4.

Reply: We have reviewed the analysis used to create Table 4 (see below). A correction in the analysis gives qualitatively the same results, but with the lags slightly reduced.

<table>
<thead>
<tr>
<th>LEVEL</th>
<th>Explained Variance</th>
<th>Correlation 150 hPa PC</th>
<th>Correlation O3 FLUX 150 hPa</th>
<th>Lag</th>
</tr>
</thead>
<tbody>
<tr>
<td>150 hPa</td>
<td>79-85%</td>
<td>NA</td>
<td>0.85 (0.43)</td>
<td>2(2)</td>
</tr>
<tr>
<td>500 hPa</td>
<td>71-77%</td>
<td>0.78</td>
<td>0.85 (0.50)</td>
<td>5(5)</td>
</tr>
<tr>
<td>1000 hPa</td>
<td>40-48%</td>
<td>0.66</td>
<td>0.73 (0.34)</td>
<td>6(6)</td>
</tr>
</tbody>
</table>

Using a scale height of 7 km, 150 hPa corresponds to approximately 13 km and 300 hPa corresponds to a height of approximately 8.5 km. We assume the tropopause is located at ~300 hPa. The average downward velocity is approximately $3.4 \times 10^{-4}$ m/s (see figure in paper). This gives a timescale of: $(13-8.5) \text{ (km)} \times 1000 \text{ (m/km)} / 3.4 \times 10^{-4} \text{ (m/s)} \sim 153$ days. Below the tropopause (e.g., ~300 hPa) the mixing is likely to be more rapid. A timescale of 5-6 months is consistent with a lower stratospheric advective timescale, roughly consistent with the lag-times obtained in the analysis.

The lag-time within the troposphere is much shorter. We obtain a lag on the order of 1 month between 500 hPa and the surface as indicated in the discussion of Figure 10.

A one-month timescale is consistent with the tropospheric lifetime of ozone. However, this timescale increases with altitude; a timescale of closer to a year is probably appropriate for the lowermost extratropical stratosphere.
We have calculated the lag for *interannual* ozone anomalies between different heights. As pointed out by the reviewer the advective timescale for ozone exchange within the troposphere can indeed be measured in days. However, the ozone anomalies associated with these individual events are of short duration and occur on small scales. The timescale for the transport of annually averaged zonally averaged ozone anomalies within the troposphere probably corresponds to a tropospheric mixing timescale. For this a timescale of a month for transport within the troposphere seems reasonable (as suggested by the discussion in association with Fig. 10). The timescale for transport from 150 hPa is much longer as suggested by the calculation above.

In addition, the correlations shown in Table 4 have a broad and smooth peak with lag-time. For example, while the correlation between the 500 hPa PC – and the 100 hPa O3 flux is 0.85 at a lag 5 months, the lag at 3 months is 0.81. Thus the results are not terribly sensitive to the lag time.

**Change:** We will include in the text that the lag is consistent with lower stratospheric advective velocities.

4. **Also,** a 3-month lag for the 150 hPa EOF correlation with ozone flux at 150 hPa seems counterintuitive. However, this could be justified if the air mass flux was significantly out of phase with the seasonality of ozone.

**Reply:** The revised calculation suggests the lag is two months. The maximum response of ozone may be expected to lag the ozone flux as the time change of ozone is proportional to the ozone flux. For example if the ozone flux corresponds to a sin-wave the ozone response would be expected to be pi/2 out of phase with it.

In addition, the ozone flux is actually calculated at 100 hPa.

**Change:** No change to this part.

**Other comments:**

5. I question why many of the figures are placed in the supplementary material. Many of these are presented and discussed on par with the rest of the material in the paper. These would not be “supplemental” and only make it more difficult on the reader to have to jump back and forth between two different places to look at the figures discussed! Most of these should be placed into the body of the paper.

**Change:** Reviewer #1 also commented on this. We will move the supplementary figures to the main body of the paper.

6. The title of the paper should reflect that the study considers only the Northern Hemisphere extratropics.
Change: We will change the title to: “Ensemble Simulations of the Role of the Stratosphere in the Attribution of Northern Extra-Tropical Tropospheric Ozone Variability” as suggested by Reviewer #1.

7. It would be helpful if it were explicitly stated earlier that these are free-running simulations. I currently see that in the beginning of the conclusion section. Also, this section could additionally describe how the ensemble members were created (differences in the initializations).

Change: We will explicitly make this point at the beginning of section 2.1. The four ensembles were initialized using different initial conditions. The revised paper will explicitly state this.

8. This really doesn’t suggest a long-term ozone decrease, especially with the large standard deviation. It appears rather flat over the long-term.

Reply: The measured record over Japan is rather spotty and noisy. However, the measured values from the late 1990s to at least 2000 are on average less than the measured values during the 1970s and early to mid 1980s. While it is true we haven’t evaluated this statistically the paper only states the data suggests a long-term ozone decrease, not a terribly strong statement.

No Change.

9. Given this statement, should the Northern Europe value in the Table be in bold?

Reply: The correlation over Northern Europe is significant at the 98% level, not the 99% significance level required to be bold in the table.

Change: We will give the significance levels in the text to remove the seeming contradiction.

10. As I understand it, the model values are averaged over the region and the observations in each region are averaged together. If so, I am not surprised that the 500 hPa measurements have a standard deviation much larger than the model but they are much more comparable at 150 hPa. The spatial variability of ozone in the troposphere is much greater than in the stratosphere. An average of a small sample of points in the troposphere (the observations) is likely to have greater variability than the average over that continental-scale tropospheric region. Figure 4 also provides supporting evidence of this. The standard deviations are quite large (and time series look completely different) during the earlier record when there are far fewer sites and measurements. After about 1980 in Canada and the early 90s in Europe, the number of sites increase and observation frequencies become greater. This corresponds to the time when the observations and model results begin to agree much better.
Reply: This is an interesting observation. However, there is likely to be more to it, at least in some locations. (1) An extensive analysis by Logan et al. (2012) shows that in fact many of the ozonesonde observations are unreliable over Europe prior to 1998. This is based partly on the correlation between individual station records. (2) The model simulations (as well as the measurements) suggest that to a large extent the annually averaged signal is highly correlated on a regional basis. While the spatial variability of the troposphere may be large, it is not as clear that the annually averaged interannual variability of the troposphere occurs on small (subregional) spatial scales.

In relation to the difference between the 500 hPa and 150 hPa records the current paper states: “This suggests a comparative degradation in the measurement accuracy at 500hPa compared to 150hPa and/or geographical variability not simulated.” This, we believe, captures the reviewer’s sense that differences between 500 and 150 hPa are related to the scale of spatial variability.

Change: We will also mention that the decrease in the standard deviation between measurement sites may relate to decreases in the number of sites and a decrease in observation frequencies.

11. P20478: It appears that the increase after 1990 could be due to the impact of Pinatubo. Thus, it appears that the 1960-2005 trend would be fairly linear if 1990-1995 were removed from the time series.

Reply: We presume the reviewer is referring to the ozone flux here. We agree that if 1990-1995 were removed from the record it may be hard to discern an increase in ozone between the late 1980 and 2005. However, there seems to be a clear decrease in the ozone flux between mid-1960s to the late 1980s (prior to Mt Pinatubo eruption).

No Change.

12. Table 4, note 3: I don’t understand exactly what you mean by “The correlation in parenthesis is computed individually for each simulation; however, the correlation coefficient comprises the overall relationship for all ensembles.”

Reply: We computed an overall correlation coefficient by correlating the overall series consisting of [S1,S2,S3,S4] with [S1*,S2*,S3*,S4*] where S1 and S1* consist of different components of the simulations (e.g., the ozone flux and the first component of the EOF).

Change: We agree that our phrasing here is rather opaque. We will change it in the revised version.

13. P20488, L6 (referencing Fig. 13): The individual titles in each panel of Figure 13 labels Mace Head, Lassen, and Alpine as the surface rather than 500 hPa. I assume it should be 500 hPa.
Reply: The measurements at the Mace Head, Lassen and the European Alpine stations are indeed measurements at the surface.

Change: The phrasing used here may have caused some confusion: “(500hPa Canadian sites, Mace Head, Lassen and the European alpine sites)”. We will change this to “(500hPa Canadian sites, and the surface sites at Mace Head, Lassen and the European Alps)”

14. P20488, L28: And the most minor comment: the words “entire” and “the” should be reversed!

Change: Thank you. We will change.

References:


Ensemble simulations of the role of the stratosphere in the attribution of northern extratropical tropospheric ozone variability

P. Hess\textsuperscript{1}, D. Kinnison\textsuperscript{2}, and Q. Tang\textsuperscript{3}

\textsuperscript{1}Cornell University, Ithaca, NY, USA
\textsuperscript{2}National Center for Atmospheric Research, Boulder, CO, USA
\textsuperscript{3}Lawrence Livermore National Laboratory, Livermore, CA, USA

Received: 19 June 2014 – Accepted: 8 July 2014 – Published:
Correspondence to: P. Hess (pgh25@cornell.edu)

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

Despite the need to understand the impact of changes in emissions and climate on tropospheric ozone, attribution of tropospheric interannual ozone variability to specific processes has proved difficult. Here we analyze the stratospheric contribution to tropospheric ozone variability and trends from 1953–2005 in the Northern Hemisphere (N. H.) mid-latitudes using four ensemble simulations of the Free Running (FR) Whole Atmosphere Community Climate Model (WACCM). The simulations are externally forced with observed time varying: (1) sea surface temperatures (SSTs), (2) greenhouse gases (GHGs), (3) ozone depleting substances (ODS), (4) Quasi-Biennial Oscillation (QBO); (5) solar variability (SV) and (6) stratospheric sulfate surface area density (SAD). Detailed representation of stratospheric chemistry is simulated including the ozone loss processes due to volcanic eruptions and polar stratospheric clouds. In the troposphere ozone production is represented by CH$_4$-NO$_x$ smog chemistry, where surface chemical emissions remain interannually constant. Despite the simplicity of its tropospheric chemistry, at many N.H. measurement locations the interannual ozone variability in the FR WACCM simulations is significantly correlated with the measured interannual variability. This suggests the importance of the external forcing applied in these simulation in driving interannual ozone variability. The variability and trend in the simulated 1953–2005 tropospheric ozone record from 30–90° N at background surface measurement sites, 500 hPa measurement sites and in the area average is largely explained on interannual timescales by changes in the 30–90° N area averaged flux of ozone across the 100 hPa surface and changes in tropospheric methane concentrations. The average sensitivity of tropospheric ozone to methane (percent change in ozone to a percent change in methane) from 30–90° N is 0.17 at 500 hPa and 0.21 at the surface; the average sensitivity of tropospheric ozone to the 100 hPa ozone flux (percent change in ozone to a percent change in the ozone flux) from 30–90° N is 0.19 at 500 hPa and 0.11 at the surface. The 30–90° N simulated downward residual velocity at 100 hPa increased by 15% between 1953 and 2005. However, the impact
of this on the 30–90° N 100 hPa ozone flux is modulated by the long-term changes in stratospheric ozone. The ozone flux decreases from 1965 to 1990 due to stratospheric ozone depletion, but increases again by approximately 7% from 1990–2005. The first empirical orthogonal function of interannual ozone variability explains from 40% (at the surface) to over 80% (at 150 hPa) of the simulated ozone interannual variability from 30–90° N. This identified mode of ozone variability shows strong stratosphere–troposphere coupling, demonstrating the importance of the stratosphere in an attribution of tropospheric ozone variability. The simulations, with no change in emissions, capture almost 50% of the measured ozone change during the 1990s at a variety of locations. This suggests that a large portion of the measured change is not due to changes in emissions, but can be traced to changes in large-scale modes of ozone variability. This emphasizes the difficulty in the attribution of ozone changes, and the importance of natural variability in understanding the trends and variability of ozone. We find little relation between the El Nino Southern Oscillation (ENSO) index and large-scale tropospheric ozone variability over the long-term record.

1 Introduction

Global change will impact both tropospheric and stratospheric ozone concentrations. Ozone acts as a potent oxidant deleterious to human health, ecosystem and agricultural productivity; it shields the Earth’s surface from harmful ultraviolet radiation; it is a greenhouse gas itself, and through its complex photochemistry regulates the lifetime of other greenhouse gases. Because of tropospheric ozone’s importance as a surface pollutant the emissions of ozone precursors are regulated in many countries while the Montreal Protocol and its amendments regulate emitted species that act to destroy stratospheric ozone. To assess the impact of emission strategies in modifying atmospheric ozone there is a need to establish clear links between measured changes in ozone and the processes that cause these changes including changes in climate and changes in emissions. Using long-term global simulations of stratospheric and tropo-
spheric ozone this paper demonstrates the large-scale coupling between extratropical tropospheric and stratospheric ozone variability. We show that an interpretation of interannual tropospheric ozone variability must account for changes in stratosphere-to-troposphere exchange (STE) of ozone.

The tropospheric ozone budget can be summarized in terms of photochemical ozone production and loss, the input of ozone from the stratosphere and the loss of ozone due to surface deposition. The largest terms, the photochemical production and loss of ozone nearly balance each other. Surface ozone deposition and influx from the stratosphere are each larger than the net photochemistry (Stevenson et al., 2006). Changes in the flux of ozone from the stratosphere to the troposphere are buffered by compensating changes in tropospheric photochemical ozone loss and surface deposition (Hess and Zbinden, 2013; Tang et al., 2013; Zeng and Pyle, 2005). Future increases in the exchange of ozone from the stratosphere to troposphere are predicted with impacts on tropospheric ozone (Stevenson et al., 2006; Collins et al., 2003; Zeng and Pyle, 2003; Shindell et al., 2006; Hegglin and Shepherd, 2009).

Except in association with particular events an overall attribution of tropospheric interannual ozone variability to specific processes has proved difficult. While very long-term ozone increases since the preindustrial are generally attributed to changes in emissions, simulations tend to underestimate the overall century time-scale ozone increases (e.g., Lamarque et al., 2005; Mickley et al., 2001) as estimated from the semi-quantitative ozone measurements at the end of the 19th century (Marenco et al., 1994; Volz and Kley, 1988). Even on the multidecadal timescales since the advent of more modern measurement techniques an attribution of measured tropospheric trends has proved difficult: the extent of the ozone increase since the 1960s as inferred from long-term N. H. measurements has not been simulated (Lamarque et al., 2010; Parrish et al., 2014).

Lin et al. (2014) attributes decadal changes in the interannual Mauna Loa ozone record to shifts in circulation patterns. However, in other locations ozone exhibits considerable interannual variability on decadal timescales that has not been adequately
explained (e.g., Koumoutsaris et al., 2008). In many cases this ozone variability is not easily ascribed to changes in emissions. For example, changes in emissions do not explain the baseline ozone trends at Mace Head, Ireland (e.g., Hess and Zbinden, 2013; Fiore et al., 2009), measured as strongly positive during the most of the 1990s but since leveling off (Carslaw, 2005; Derwent et al., 2007; Simmonds et al., 2004; Derwent et al., 2013). In an analysis of ozone trends over Europe Wilson et al. (2012) conclude the impact of European precursor emission reductions were masked by other sources of unknown ozone variability. Analysis by Logan et al. (2012) and Cui et al. (2011) show the measured ozone increases at alpine sites over Europe during the 1990s followed by decreases after 2000 are not easily explained by changes in emissions or changes in lower stratospheric ozone. Pozzoli et al. (2011) conclude that changes in meteorology and natural emissions account for 75% of ozone variability from 1980–2005, largely masking changes in anthropogenic emissions. On decadal timescales ozone trends can depend sensitively on the exact time-period examined (Cui et al., 2011).

While an overall attribution and synthesis of tropospheric ozone variability may be lacking, sources of variability associated with various events have been isolated. The analysis of Leibensperger et al. (2008) shows that decadal and interannual cyclone trends have important impacts on surface ozone variability and trends over the North East US, although the analysis of Turner et al. (2013) suggests cyclones explain less than 10% of the overall variability of high ozone events over this region. Over the Eastern half of the US changes in interannual temperatures explain between approximately 25–50% of the interannual regional ozone variability in July, although changes in temperature mask many other processes. Jaffe et al. (2008) associates interannual variability in ozone across the Western US with biomass burning. Interannual variability in ozone due to heat waves (sometimes also associated with changes in emissions) has also been shown to be important (Konovalov et al., 2011; Tressol et al., 2008).

It is well known that stratospheric ozone can be transported into the troposphere (e.g., Danielsen, 1968). Vertical correlations between stratospheric and tropospheric ozone have been analyzed in measurements (Hess and Zbinden, 2013; Tarasick, 2005; ...
Thouret et al., 2006; Ordóñez et al., 2007, Neu et al. 2014) and in a model simulations (Terao et al., 2008; Hess and Zbinden, 2013). A number of modeling studies have attributed extratropical N. H. tropospheric ozone variability to El Nino Southern Oscillation (ENSO) (Zeng and Pyle, 2005; Doherty et al., 2006; Koumoutsaris et al., 2008; Voulgarakis et al., 2011) modulated through STE (Zeng and Pyle, 2005; Voulgarakis et al., 2011). Langford et al. (1998) attribute modulation of middle and upper tropospheric ozone to ENSO using LIDAR measurements over Colorado. They suggest that this modulation may induce different long-term decadal ozone trends (between -0.2 to +0.5 ppbv yr⁻¹) depending on the exact period the ozone trend is examined. On the other hand, Hsu and Prather (2009) do not find a relationship between ENSO and STE. Other studies have attributed ozone variability to the North Atlantic Oscillation (NAO) or Artic Oscillation (AO) (Li, 2002; Creilson et al., 2003, 2005; Lamarque and Hess, 2003; Hess and Lamarque, 2007; Sprenger, 2003; Pausata et al., 2012) with associated changes in STE (Hess and Lamarque, 2007; Sprenger, 2003). Hsu and Prather (2009) show considerable interannual variability in stratosphere–troposphere exchange and attribute 20–40 % of this variability to the quasi-biennial oscillation (QBO). Tropospheric ozone decreases have been simulated following the Mt Pinatubo eruption due to changes in STE (Tang et al., 2013). Hess and Zbinden (2013) argue that to a significant extent interannual variability in extratropical tropospheric ozone is due to the variability in ozone transported from stratosphere. Neu et al. (2014) attribute approximately half of tropospheric ozone variability to interannual changes in the strength of the stratospheric circulation.

In this paper we use a synthesis of simulations and measurements to demonstrate the importance of large-scale coupled stratosphere–troposphere modes in determining tropospheric ozone variability from 30–90° N. These results are an extension and expansion of the simulations analyzed in Hess and Zbinden (2013), who showed the importance of stratosphere–troposphere exchange in explaining N. H. extratropical tropospheric ozone variability from 1990–2006. We expand on the work of Hess and Zbinden (2013) by: (1) using simulations with good stratospheric resolution and...
Detailed representation of stratospheric chemistry incorporating the impacts of interannual changes in stratospheric aerosol loading and ozone depleting substances (ODS).

(2) Simulating the chemical coupling between the stratosphere and troposphere over a period of more than 50 years (1953–2005), a period incorporating the rapid growth and then decline of the emissions of ODS. (3) Analyzing the extent to which the coupled variability of the lower stratosphere and tropospheric ozone is externally (e.g., by changes in sea surface temperatures) vs. internally forced. (4) Incorporating further analysis of the large scale coupled modes linking stratospheric and tropospheric ozone variability.

Distinct from Hess and Zbinden (2013) we use a simulation that only simulates basic tropospheric NOx-CH4 chemistry. By examining the importance of stratospheric-tropospheric coupling using a basic set of tropospheric chemistry reactions, the importance of more complex chemistry in determining tropospheric ozone variability can be better understood. It is expected that the introduction of additional hydrocarbon chemistry as well as episodic emission variability (e.g., biomass burning) will introduce additional modes of variability not captured here. In addition more complex chemistry may possibly dampen the basic modes of ozone variability described below. However, despite the simplicity in the tropospheric chemistry, these simulations match the observed variability to a large extent. Thus we view the modes of ozone variability captured here as base-state modes which may be perturbed by more complex chemistry, but are fundamental to the coupled troposphere-stratosphere chemical system.

The model description and description of the data analyzed is given in Sect. 2. An evaluation of the simulations is given in Sect. 3. Section 4 analyzes the ozone variability. Discussion and conclusions are given in Sect. 5.
2 Methodology

2.1 Model description

The Whole-Atmosphere Community Climate Model, Version 3.5 (WACCM3) is a comprehensive numerical model simulating the dynamics and chemistry of the atmosphere, spanning the range of altitude from the Earth’s surface to the lower thermosphere. Four long-term ensemble free running (FR) WACCM simulations for the period from 1953–2005 are generated by initializing the model with different initial conditions. The analysis of these simulations focuses on the tropospheric and lower stratospheric ozone distribution from 30–90° N.

WACCM3 is a fully interactive model, wherein the radiatively active gases (CO₂, H₂O, N₂O, CH₄, CFC-11, CFC-12, NO, O₃) affect heating and cooling rates and therefore dynamics (Sassi et al., 2005). WACCM is based on the software framework of the Community Atmospheric Model (CAM). WACCM3, is a superset of CAM version 3 (CAM3), and includes all of the physical parameterizations of that model. A finite volume dynamical core (Lin, 2004), which is an option in CAM3, is used exclusively in WACCM3. This numerical method calculates explicitly the mass fluxes in and out of a given model grid cell, thus ensuring mass conservation.

The governing equations, physical parameterizations and numerical algorithms used in CAM3 are documented by Collins et al. (2006); only the gravity wave drag and vertical diffusion parameterizations are modified for WACCM3. In addition, WACCM3 incorporates a detailed neutral chemistry model for the middle atmosphere, including heating due to chemical reactions; a model of ion chemistry in the mesosphere/lower thermosphere (M/LT); ion drag and auroral processes; and parameterizations of short wave heating at extreme ultraviolet (EUV) wavelengths and infrared transfer under non-local thermodynamic equilibrium (NLTE) conditions. Processes and parameterizations that are unique to WACCM3 are discussed in Garcia et al. (2007); for other details, the reader is referred to the papers of Collins et al. (2006).
The chemistry module is based on the Model for OZone And Related chemical Tracers version 3 (MOZART3) (Kinnison et al., 2007). The species included within this mechanism are contained within the O\(_x\), NO\(_x\), HO\(_x\), ClO\(_x\), and BrO\(_x\) chemical families, along with CH\(_4\) and its degradation products (a total of 59 species and 217 gas-phase chemical reactions). This chemical mechanism includes 10 long-lived organic halogens (i.e., CH\(_3\)Cl, CFC-11, CFC-12, CFC-113, HCFC-22, CH\(_3\)CCl\(_3\), halon-1211, halon-1301, and CH\(_3\)Br). Rate constants are based on Sander et al. (2006). In addition, there are 17 heterogeneous reactions on three aerosol types: Nitric Acid Trihydrate (NAT), Supercooled Ternary Solution (STS), and Water-Ice. A detailed description of the chemical approach can be found in Kinnison et al. (2007).

For this work, the Chemistry Climate Model Validation Activity for SPARC, version 2 (CCMVal2) REF1 scenario was used (see Eyring et al., 2008). This scenario included observed time-dependent evolution of: greenhouse gases (GHGs); ozone-depleting substances (ODSs); sea surface temperatures and sea ice concentrations (SSTs/SICs); stratospheric sulfate surface area densities (SADs); 11 year solar cycle variability, which includes spectrally resolved solar irradiances; Quasi-Biennial Oscillation (QBO), by relaxing to observed tropical winds. Surface emissions of CO, NO\(_x\), and formaldehyde are included but the emission trends are not simulated. The emissions are set to present day conditions. This version of WACCM was extensively evaluated in the SPARC Report of the Evaluation of Chemistry-Climate Models (SPARC-CCMVal, 2010).

Figure 1 gives the change in boundary conditions for CH\(_4\) and CFC-11. CFC-11 peaks in 1992 (World Meteorological Organization (WMO), 2007) and then begins to slowly decline. Methane shows a nonlinear growth rate with evidence of a flattening trend beginning in the early 1990s (e.g., Dlugokencky et al., 2011).
2.2 Data

The WACCM simulation is evaluated at 150 hPa, 500 hPa, and a number of surface stations between 30 and 90° N including elevated alpine sites. Long-term monitoring sites and ozonesonde records are used for model evaluation (see Table 1).

Ozonesonde data is obtained from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC). The methodology for filtering the measurements is similar to that described in Hess and Zbinden (2013) except we combined different measurement techniques together so as to extend the measurement records as far back as possible. As in Hess and Zbinden (2013) we aggregated the 150 and 500 hPa ozonesonde data into geographical regions. This acts to isolate the larger-scale interannual variability and increases the sampling frequency. We use geographical regions located between 30 and 90° N where at least two long-term independent ozonesonde measurements are available: Canada, Central Europe, Japan and Northern Europe (see Table 1). Here we simply aggregate the regional ozone measurements by averaging the individual measurements within each region. Hess and Zbinden (2013) aggregated the measurements by averaging their relative variability, but the two methodologies produce very similar results.

At the surface, we include many of the same long-term measurement sites between 30 and 90° N as used in Lamarque et al. (2010) (see Table 1). We have combined the Zugspitze alpine measurements with those of the neighboring Jungfraujoch surface measurement site. We have omitted measurement sites immediately downwind of Asian or US emissions (Mt Happo, Japan; Bermuda; Sable Island, Nova Scotia) as our simulations are best suited to sampling background air as we include no changes in surface emissions. We have, however, included the Lassen National Park site in the Western US even though this site likely registers impacts of increasing Asian emissions (e.g., Cooper et al., 2010; Parrish et al., 2012). This site is subject to significant interannual variability not explained by changes in Asian emissions (see Fig. 5). We
have also omitted the Barrow site due to the influence of Arctic depletion events on the Barrow record (e.g., Oltmans et al., 2012).

For each measurement site, or measurement region, monthly ozone deviations are calculated as deviations from the monthly-averaged ozone distribution from January 1990–December 2004. The monthly deviations are averaged using 12 month smoothing.

3 Model evaluation

3.1 Simulation overview

Ozone for the four ensembles simulations averaged from 30–90° N is shown in Fig. 2 at 150, 500 and the surface. At each level for each of the ensemble simulations ozone follows a similar long-term trajectory. At 150 hPa this trajectory is at least partially governed by the ODS forcing common to all ensemble members. At 150 hPa, over the long-term, ozone levels remain nearly constant until 1970, but thereafter decrease and reach a minimum between 1991 and 1994 in association with the Mt Pinatubo eruption. Note that even though all ensemble members simulate ozone loss associated with Mt Pinatubo, the timing of this loss is partially governed by the internal model variability. In all simulations 150 hPa ozone partially recovers following the Mt Pinatubo period. At 500 hPa and the surface all simulations show a long-term ozone increase throughout the period. This increase can be largely attributed to the long-term methane forcing (Fig. 1). The long-term growth in ozone is approximately linear until 1980 when it begins to flatten.

On shorter timescales each of the ensemble simulations follows its own path relative to the mean long-term trend. However, there are some notable periods where all ensemble members show similar behavior with a coincident ozone maximum or minimum across all model levels (e.g., 1965, 1967, 1972, 1973, 1998); during other periods the short timescale behavior of the individual ensemble members appear unrelated (e.g.,
1968–1971, 1977–1983, 1993–1997). In the sections below the model behavior depicted in Fig. 2 is examined in detail. In this section the model measurement evaluation is given at selected sites from 150 hPa to the surface (see Table 1).

3.2 Evaluation

The model-measurement evaluation is shown Figs. 3–6 and is summarized in Table 2. A comparison of the area averaged ozone evolution (Fig. 2) and that at the individual ozone sites reveals broad similarities. The similarity between the ozone evolution at various measurement sites and the overall hemispheric evolution was pointed out in Hess and Zbinden (2013). This is further examined in Sect. 4 where we examine coupled stratospheric-tropospheric modes of ozone variability.

The model-measurement correlations (Table 2) are between 12 month smoothed datasets and use detrended data so as not to introduce positive correlations simply through a simulation of long-term trends. In the stratosphere ozone is simply detrended against time; in the troposphere ozone is detrended against methane as changes in methane largely drive the simulated ozone increases (see Sect. 4). We note it is difficult to unambiguously separate out the shorter term interannual variations from the longer term trends, particularly when the long-term trends have a comparable magnitude to the interannual variations. We believe this is particularly problematic in the analysis of some of the tropospheric measurements with large ozone increases (e.g., at Lassen). The modeled ozone is obtained from monthly averaged model output at all measurement stations within a region. No attempt is made to duplicate the temporal sampling of the measurements at each measurement site. The model-measurement correlation is broken into two periods: prior to 1990 and subsequent to 1990. The larger number of measurement sites and the increase in measurement accuracy make comparisons after 1990 particularly valuable. The model and measured record are only correlated when the records extend for periods of greater than five years. Shorter-term model-measurement comparisons (1991–1995) are made in Tang et al. (2013) using a similar formulation of WACCM as used here, but with full hydrocarbon chemistry in the
troposphere. Tang et al. (2013) also compare the measurements against simulations using a specified dynamics formulation of WACCM, where WACCM is nudged to analyzed winds. The more tropospherically configured model, the Community Atmospheric Model with chemistry (CAM-chem) (Lamarque et al., 2012) has been compared against a similar set of measurements in Hess and Zbinden (2013) and Tang et al. (2013).

Not until the early 1970s do ozonesonde measurements become available for the Canadian, Central European and Japanese sites (Table 2). Data is not available for the Northern European Region until the late 1980s. Figures 3–4 give the observational record for each ozonesonde site within each region and document the changes in the type of ozonesonde used at each site. The accuracy of the regionally averaged ozone records likely change with time as the number of stations and measurement techniques change (e.g., from Brewer Mast (BM) ozonesondes to electrochemical concentration cells (ECC)). The standard deviation between the regional measurements shown at the bottom of each figure gives an indication of temporal changes in the regional consistency of the measurements.

### 3.2.1 150 hPa evaluation

At 150 hPa simulated ozone remains fairly flat at the analyzed sites (e.g., Fig. 3) until the early 1970s when the earliest measurements become available. Coincident with the increase in the concentrations of ODS (e.g., see Fig. 1) simulated ozone decreases from the 1970s until the early 1990s over the four regions examined (Canada, Central Europe, Japan and Northern Europe). Measured decreases during this period are particularly notable over the Canadian and Central European regions. While the standard deviation between the regional measurement sites is comparatively large prior to 1990 over the Canadian region (Fig. 3), the standard deviation is still smaller than the overall long-term trend (1970–1990). The early record over Japan is somewhat noisier, but also suggests long-term ozone decreases during this period (Fig. 3). The negative ozone deviations at 150 hPa in the early 1990s can be attributed to Mt Pinatubo, which resulted in significant ozone depletion in the northern mid-latitudes beginning in 1991.
and lasting through 1995 (e.g. Tang et al., 2013). Since the early 1990s the 150 hPa ozone at the evaluated sites has recovered to some extent.

In addition to the long-term trends the simulations and measurements also exhibit considerable shorter-term interannual variability at 150 hPa. For the period beginning in 1990 the model and measured records are significantly correlated over Canada (at the 99 % level) and Northern Europe (at the 98 % level) (Figs. 3 and Table 2). For example after 1990 both measurements and simulated ozone over Canada and Northern Europe have an ozone minimum during the Mt Pinatubo period, a maximum in 1998–1999 and 2002–2003 and pronounced ozone minima near 2000. These features can also be seen in the Central European record (Fig. 3). Prior to 1990 the detrended simulated and measured variability are significantly correlated over Canada and Central Europe (Table 2) at 150 hPa. Over Canada the standard deviation between the individual station measurements is relatively large prior to 1990, suggesting caution in interpreting the early measurements. Nevertheless, the peak in the Canadian measurements in 1970 and 1973 and the broad peak from 1977–1983 correspond to similar features in the Central European measurements suggesting that these features may be real. The simulation and measurements are not significantly correlated over Japan. Hess and Zbinden (2013) found the measurements from the individual stations over Japan are not temporally coherent at 150 hPa. This is evident from the rather high standard deviation between the ozonesonde stations (Fig. 3). In fact we have not been able to simulate the ozone record over Japan either prior to 1990 or subsequent to 1990 at either 150 hPa or at 500 hPa in the current simulations.

3.2.2 500 hPa evaluation

While the 150 hPa simulated and measured standard deviation between regional sites was similar in the model and measurements (Figs. 3), at 500 hPa the measured standard deviation is much larger than that simulated (Figs. 4). This suggests a comparative degradation in the measurement accuracy at 500 hPa compared to 150 hPa and/or geographical variability not simulated. In particular, the decrease in
the number of regional measurement sites during the early periods may accentuate the geographic differences in the measured ozone, particularly in the troposphere where the spatial scale of variability is likely to be smaller than in the stratosphere. Over Europe in particular, the measured standard deviation increases significantly prior to 1990. The analysis of Logan et al. (2012) shows the ozonesonde data has only been coherent over Europe since 1998. Hess and Zbinden (2013) also noted discrepancies in the European data during the 1990s.

Over the long-term significant ozone increases are both simulated and measured over the 500 hPa tropospheric sites. In the simulation these increases can be attributed to the long-term increases in methane (Fig. 1); in the measurements, the emissions of many other ozone precursors also increased over this time period. Prior to the early 1980s simulated ozone significantly over predicts the tropospheric measurements, except over Japan. We note that even simulations using complex mechanisms for tropospheric photochemistry and time varying emission inventories have not been able to capture the measured ozone increases since 1980 (Lamarque et al., 2010; Parrish et al., 2014). The Canadian sites show a rapid ozone increase until 1980, while the Central European sites show this increase lasts until the middle to late 1980s, consistent with European Alpine sites (see Fig. 5). However, whereas the Canadian and alpine sites suggest ozone in the mid-1980s was comparable to the concentrations after 2000, the Central European ozonesondes show much elevated concentrations during this earlier period. The measured high ozone concentrations over Central Europe at 500 hPa are not simulated.

Except over Japan the model and measurements are significantly correlated after 1990 at 500 hPa over the evaluated regions (Table 2). During the early 1990s the eruption of Mt Pinatubo resulted in anomalously low tropospheric ozone (Tang et al., 2013) clearly evident over the Central and Northern European and Canadian regions in both the simulations and measurements (Figs. 4 and S4). The elevated ozone in 1998–1999 evident in the measurements and simulations (also see Thouret et al., 2006; Zbinden et al., 2006) has been attributed to ENSO (Koumoutsaris et al., 2008; Voulgarakis et al.,
2011) and has been associated with increases in STE (Voulgarakis et al., 2011; Hess and Zbinden 2013). The model and measurements are also significantly correlated over the Canadian ozonesonde sites prior to 1990.

### 3.2.3 Surface evaluation

As at 500 hPa, the European Mountain sites (Fig. 5a), the Mt Lassen site in California (Fig. 5b), the Mace Head site in Western Europe (Fig. 6a) and the Arkona site in Continental Europe (Fig. 6b) all exhibit an ozone minimum in the measurements and simulations near 1993–1994 (although not very pronounced in the Mace Head measurements) and an ozone maximum near 1998–1999. The simulations and measurements are significantly correlated at the European mountain sites after 1990 (Fig. 5a) and at Mace Head (Fig. 6a), but not at Lassen (Fig. 5b) or Arkona (Fig. 6b). The ozone record at Lassen is particularly susceptible to the large increase in Asian ozone precursor emissions (e.g., Cooper et al., 2010; Parrish et al., 2012), increases not included in the simulation. Interannual variability in the transport of ozone produced from these Asian emissions as well difficulties in unambiguously removing the ozone trend at Lassen may contribute to the low model-measurement correlation at that site. The model simulations, assuming no increases in emissions capture much of the measured variability and ozone change that occurred since 1990 at these sites. The Arkona site is situated over continental Europe in a region immediately impacted by European emissions. At this site the simulations dramatically over estimate the measured concentrations prior to 1990 (Fig. 6b) and do not capture the ozone variability subsequent to 1990.

### 4 Long-term tropospheric ozone variability

#### 4.1 Forced vs. unforced variability

Over the long-term the trends in the simulated ozone are driven by the trends in the concentrations of ODSs and methane and the solar cycle. Short-term external forcing
can be attributed to forcing by sea surface temperature, volcanoes and the QBO. To the extent that the ozone record is driven by internal model dynamics versus interannually varying external forcing we would expect the ozone records from the different ensemble members to be uncorrelated with each other and uncorrelated with the measurements. Given a perfect model (and perfect measurements) the correlation between simulations and measurements should give an indication of the importance of external forcing to the simulations. The positive and significant model-measurement correlations at various sites, particularly for the period after 1990 (see Table 2), in simulations in which model dynamics is internally calculated, emphasizes the importance of forced variability in driving the ozone variations.

The correlation between ensemble members also provides an indication of the extent to which the model is externally forced. The average median detrended correlation between the different simulations at all sites is highly significant (see Table 2), suggesting the role of external forcing is significant. These correlations tend to be somewhat lower at the surface. These lower correlations likely reflect the shorter lifetime of ozone near the surface and thus an increased importance of local and regional processes. The correlations are lowest at Arkona suggesting a decreased role for forced variability in association with high surface emissions. The ensemble correlations suggest that between 6 and 19 % of the ozone variability at the surface measurement sites can be explained by external forcing and between 16 and 25 % of the variability is forced at 500 and 150 hPa. The ensemble correlations over the course of the simulation are generally less than the model-measurement correlations, but the high model-measurement correlations generally occur after 1990 when the external forcing due to Mt Pinatubo and the 1998–1999 El Nino is particularly strong.

4.2 Response of tropospheric ozone to changes in methane and stratospheric ozone flux

In addition to photochemical changes in the tropospheric ozone budget, tropospheric ozone is modulated by the flux of ozone from the stratosphere (e.g., Stevenson et al.,
The vertical component of the zonally averaged vertical residual velocity \((\bar{w}^v)\) and the ozone flux across 100 hPa averaged from 30–90° N is shown in Fig. 7 for each ensemble member. The vertical component of the residual circulation \((\bar{w}^z)\) approximates the vertical transport circulation (Andrews et al., 1987) and serves as a measure of the strength of the overturning Brewer and Dobson circulation. The averaged downward residual velocity at 100 hPa increases from 1953–2005 (Fig. 7a) in all the ensemble simulations. We note during some years \(\bar{w}^v\) in each of the ensemble simulations is similar, while during other years \(\bar{w}^v\) appears rather unrelated between the different ensemble members. On average the residual circulation increases downward by approximately 15% over the course of the simulation. This increase in the residual circulation is a robust feature of chemistry climate model simulations (Butchart et al., 2006, 2010; Eyring et al., 2010; SPARC-CCMVal, 2010; Oman et al., 2010). The simulated increase in the downwards component of \((\bar{w}^z)\) averaged from 30–90° N (Fig. 7a) is due to both an increase in the strength of the downward component of \((\bar{w}^z)\) and an increase in the area in which \((\bar{w}^z)\) is downwards. The latter is consistent with a narrowing of the upwelling branch of the Brewer-Dobson circulation in response to climate change (Li et al., 2010; Hardiman et al., 2014). A cubic fit (Fig. 7) suggests this increase is not exactly linear, but has increased since 1990.

The 100 hPa ozone flux averaged from 30–90° N is calculated by multiplying the residual vertical velocity by the ozone concentration. Gettleman et al. (1997) suggest that the flux of ozone across this level serves as a good proxy for the flux of ozone from the stratosphere to the troposphere. While diagnostics of the STE of ozone across the tropopause would be preferable, they could not be estimated precisely from the monthly averaged model output fields saved from these simulations. The ozone flux from 30–90° N averages 352 Tg/year while the global flux averages 513 Tg/year. The global flux is similar to the estimate in Hsu and Prather (2009). In contrast to the fitted increase in the downwards-residual velocity (Fig. 7a), the change in the ozone flux has not been monotonic (Fig. 7b). The cubic fit to the ozone flux reaches a maximum in the 1960s as the residual circulation increases in strength and stratospheric ozone
remains fairly constant; following this period the ozone flux decreases until the early 1990s (corresponding to the time of the Mt. Pinatubo eruption) as stratospheric ozone decreases; following Mt Pinatubo the ozone flux increases again until the end of the model simulation. On average the ozone flux has increased by 8 % from 1953–2005, about half the rate of the increase of the residual circulation. Hegglin and Shepherd (2009) also show the ozone flux is modulated by ozone depletion in the N. H., but suggest no long-term decrease in the flux with the smallest fluxes occurring near 2000. We note the simulations described in Hegglin and Shepherd (2009) include no forcing due to volcanoes. In the future, predicted stratospheric ozone recovery and predicted increases in the strength of the residual circulation are expected to lead to further increases in the stratosphere-to-troposphere exchange of ozone (e.g., Hegglin and Shepherd, 2009).

The long-term impact of changes in the stratospheric ozone flux on tropospheric ozone is clearly seen in Fig. 8. Here we examine the evolution of normalized ozone against tropospheric methane. Ozone is normalized by dividing the ozone record by the average ozone concentration from 1980–1985: Parrish et al. (2014) found that normalizing ozone helped reduce the ozone record at different measurement sites to a common curve. The exact date used for normalization is arbitrary. We found the data displayed nicely when we used the ozone average value from 1980–1985 to normalize the ozone record. The simulated record of normalized ozone plotted against globally averaged methane reduces the ozone record to a set of similar curves at various tropospheric measurement sites (Fig. 8a). We find that a cubic fit of ozone against tropospheric methane captures some of the subtleties of the long-term record better than a quadratic fit. Discussion of various polynomial fits is given in Parrish et al. (2014). Here we only display those locations (see Table 1) where the ozone measurements extend to the early 1980s.

At all sites, the simulated ozone increase with respect to methane is relatively fast between approximately 1953–1970 and 1990–2005, but is noticeably slower from 1970 to 1990. As the simulated surface NO$_x$ emissions are not increasing over the course
of the simulation, we would expect the ozone production from the simulated methane NO\textsubscript{x} tropospheric chemistry to be approximately linear in methane. The simulated long-term non-linearity in ozone with respect to methane is consistent with the long-term modulations in the flux of ozone from the stratosphere: the period of slow tropospheric ozone growth (from approximately 1970–1990) is coincident with the decrease in the flux of ozone from the stratosphere.

It is difficult to draw any strong conclusions through a comparison of the simulated and measured ozone-methane curves (Fig. 8). The measured curves show a variety of different shapes. The measured curve over Central Europe reaches a maximum near 1990 and then decreases. This curve is undoubtedly impacted by the high measured ozone concentrations over Central Europe in the 1980s (Fig. 4). These early measurements should be treated with caution. The Japanese measurements also suggest an early ozone maximum (in the 1980s), but show an ozone increase in the latest period. Other measured curves show an ozone increase near the end of the period. The cubic fit to the data is likely sensitive to the length of the data record as well as to the beginning and end points of the measured timeseries.

In the measured curves it is difficult, if not impossible, to separate out the impact of changes in the emissions of ozone precursors from that due to changes in the influx of ozone from the stratosphere. If, for example, we concentrate on changes in the emissions of NO\textsubscript{x}, the emissions increased rapidly in Europe from 1950–1980 by approximately 500% (Vestreng et al., 2009), flattened out between 1980 and 1990 and decreased thereafter. The US emissions stabilized following the clean air act of 1970 (EPA, 2000) after increasing by approximately 250% from 1950–1970. Emissions over East Asia increased by approximately 250% between 1980 and 2000, but were not commensurate with either US or European emissions until approximately 1995 (Ohara et al., 2007). Since the simulation does not include these changes in ozone precursor emissions it is not surprising that measured ozone changes (Fig. 8) are about a factor of three to four larger than those simulated. The measured sharp increase in ozone prior to 1970–1980 and then the transition to a flatter trend is consistent with the emis-
After 1985 the shape of the various measured curves is not consistent with the emission trends.

The coefficients for the multiple regression of simulated normalized ozone against normalized methane and the normalized stratospheric 100 hPa ozone flux (averaged from 30–90° N) are given in Table 3 at various measurement sites and for the 30–90° N area-averaged ozone. The normalized fields are obtained by dividing the respective fields by their average value from 1980–1985. In this regression we average across the four ensemble members. The best fit is obtained when the tropospheric response lags the ozone flux by approximately 5–6 months. This time lag is consistent with the relatively small lower stratospheric residual velocities (Fig. 7a). It is also roughly consistent with Liang et al. (2009) and discussion therein that it takes one month for a tracer to cross the tropopause from the lower stratosphere, another month to descend to the mid-troposphere and an additional month to reach the lower troposphere. In our study we are including the additional transport time from 150 hPa to the tropopause. At all sites correlations between the regressed fit and the simulated ozone are highly significant and greater than 0.9 (see Fig. 9). The sensitivity of normalized tropospheric ozone (i.e., the fractional change in tropospheric ozone) to fractional changes in methane and to fractional changes in the 100 hPa ozone flux is roughly similar, about 10–20 % (Table 3). The sensitivity to the ozone flux is generally higher at 500 hPa than at the surface, with the largest sensitivity at the northernmost 500 hPa ozonesonde sites (0.24–0.25). The sensitivity to the methane is higher at the surface than at 500 hPa with the highest sensitivity at Arkona (0.45). The high sensitivity to methane at Arkona is likely due to the locally high NOx emissions at the site. We also note that the regressed fit is poorest at Arkona. The sensitivity of the overall 30–90° averaged ozone concentration is similar to the sensitivity at the various sites. The sensitivity coefficients will likely be impacted by the tropospheric chemical mechanism, although the impact may not be large under background chemical conditions.

As discussed above the non-linearity in the simulated long-term tropospheric ozone trend with respect to methane can be largely ascribed to the long-term modulation of
the ozone flux at 100 hPa. However, the regression also captures many of the short-term changes in the simulated ozone (Fig. 9). Since methane is only slowly changing, the short term ozone variability is due to variations in the flux of ozone across 100 hPa. Interestingly, the regressed fit does not capture the ensemble average ozone change during the Mt Pinatubo period. The portion of the simulated ozone record not explained by methane changes is obtained by subtracting the dependence of ozone on methane (determined from the regressed fit) from the simulated ozone record. In Table 3 we give the correlation between this quantity and the vertical flux of ozone across 100 hPa averaged from 30°–90° N. The area-averaged correlations are large, ranging from 0.84 at 500 hPa to 0.73 at the surface. At 500 hPa the correlation at individual sites is similar to the area-wide correlation except over Japan, although even over Japan the correlation is significant. At the surface the correlations at elevated sites are similar to the area-wide correlation. The variability explained by the stratospheric flux is low and not significant at Arkona, situated in a region of relatively large local emissions. At Mace Head the stratospheric flux is only marginally correlated with the simulated record (Table 3). The Mace Head site is significantly impacted by European emissions. Sampling the simulations northwest of Mace Head (by 10° longitude west and by 5° latitude north) significantly improves the correlation with the stratospheric ozone flux (Table 3). In the absence of daily output data from these simulations, it is likely that this displaced location will be more representative of the filtered baseline ozone measurements at Mace Head (Derwent et al., 2007; Simmonds et al., 2004) than sampling the model at the actual location of Mace Head. Hess and Zbinden (2013) found that the stratospherically tagged ozone has a large influence on the variability at the Mace Head site.

4.3 Response of tropospheric ozone to stratospheric ozone perturbations

The large-scale area-averaged simulated ozone variability is highly correlated across vertical levels (Fig. 10) in all simulations. Nue et al. (2014) shows the correlation between stratospheric and tropospheric ozone is a good proxy for the relationship between tropospheric ozone and changes in STE. In this figure we subtract out the
linear growth of tropospheric ozone due to methane changes. Stratospheric ozone is not detrended in this analysis so the impact of stratospheric ozone depletion and recovery on long timescales is retained. Detrended tropospheric ozone at the 500 and 1000 hPa levels is consistent with the long-term trends in stratospheric ozone: at both levels detrended tropospheric ozone is a maximum near 1970 and a minimum during the period influenced by the Mt Pinatubo eruption in the early 1990s. The shorter-term year-to-year fluctuations in ozone are also highly correlated in all simulations for each ensemble member. The correlation between the 150 hPa area averaged ozone and the detrended tropospheric area average at 500 hPa reaches 0.80 with a lag of 3 months; the correlation between 150 hPa and surface ozone reaches 0.75 with a lag of 4 months; the correlation between area-averaged 500 hPa ozone and the surface reaches 0.90 with a lag of 1 month. All these correlations are highly significant. Stratospheric-tropospheric correlations based on satellite data between are smaller: Neu et al. (2014) find the correlation between the 40–50° N 150 hPa ozone anomaly and that at 510 hPa between 30–50° is 0.40. Hess and Zbinden (2013) also discussed large-scale modeled and measured ozone correlations between the lower stratosphere and the surface, and found the correlations to be significant. On a more regional or local scale Tarasick (2005), Ordóñez et al. (2007), Thouret et al. (2006) and Terao et al. (2008) have reported significant measured correlations between stratospheric and tropospheric ozone. Neu et al. (2014) shows vertical correlation in satellite retrieved ozone anomalies between 150 and 500 hPa. Tang et al. (2013) simulates the impact on tropospheric ozone of the large-scale stratospheric ozone reductions during the Mt Pinatubo period.

The overall response of 500 hPa ozone (averaged from 30–90° N) to changes in 150 hPa ozone (averaged from 30–90° N) is 0.018 ppb ppb⁻¹ (not shown). Given a 150 ppb decrease in stratospheric ozone between approximately 1970 and the Mt Pinatubo period, the resulting ozone at 500 hPa would have decreased by 2.7 ppb as a result. At individual stations the tropospheric response is stronger. The slope of surface ozone to stratospheric perturbations is 0.007 ppb ppb⁻¹. This implies an area wide
ozone decrease of approximately 1 ppb due to stratospheric ozone depletion. The overall response at the surface to ozone perturbations at 500 hPa is 0.38 ppb ppb⁻¹. As shown by Hess and Zbinden (2013), Tang et al. (2013) and Zeng and Pyle (2005) the tropospheric ozone response to increased STE is buffered by increases in tropospheric chemical ozone loss and deposition.

4.4 Coupled modes of variability

The spatial pattern of ozone variability at 150 hPa, 500 hPa and the surface is analyzed using empirical orthogonal function (EOF) analysis. This analysis separates ozone variability into orthogonal basis functions. Each function is specified by a spatial pattern (with a dependence on location only) and a time-series for the temporal variations of this pattern with a dependence only on time (the principal component time series). EOF analysis allows an understanding of the geographic variability of the ozone record and relates the variability at different locations.

The EOF analysis is conducted on the detrended ozone record at all locations from 30–90° N. The ozone record at the surface and 500 hPa are detrended by regressing ozone against global methane concentrations; the ozone record at 150 hPa is detrended with respect to time. The first EOF component at all three levels is given in Fig. 11. The EOF is normalized so that its value gives the standard deviation of the ozone variations due to this EOF; the sign specifies the phase difference between the ozone variations explained by the EOF. Points with different sign have opposite temporal phases. The EOF captures from 40–48 % of the ozone variability at the surface, 71–77 % of the variability at 500 hPa and 79–85 % of the variability at 150 hPa (Table 4).

For each ensemble member the correlation between the temporal variability of the principal component and the detrended area average ozone is very high (greater than 0.95 on all three levels, not shown). Thus, the temporal variation in EOFs is closely related to the area averaged ozone variations. However, the use of EOFs refines the simple use of area-averages by showing geographical differences in the pattern of variability with a better statistical characterization of the variability.
For each ensemble member, the principal component timeseries are highly correlated across vertical levels (Fig. 13 and Table 4) suggesting the modes of variability isolated by the EOF analysis are physically deep (i.e., they span from at least the surface to 150 hPa). Due to the large ozone gradients between the stratosphere and troposphere, it is difficult to escape the conclusion that the coupled variability between the stratospheric and tropospheric levels is linked through the transport of high stratospheric ozone concentrations to the troposphere (see Neu et al., 2014). This is consistent with the analysis of Hess and Zbinden (2013). As discussed below, the geographical pattern of the variability supports this conclusion.

The ensemble average of the area-averaged 30–90° N ozone flux at 100 hPa explains 40 % of the variability of the ensemble average principal component timeseries at the surface, 58 % at 500 hPa and 69 % at 150 hPa (with a lag of 3 to 9 months) (Table 4). The lag increases as one descends in the atmosphere consistent with timescales for the transport of ozone from the lowermost stratosphere to the surface. The correlation between the ozone flux for each ensemble member and the principal component time series for that ensemble member (instead of the correlation between the ensemble averages) reduces the variability explained by the ozone flux to between approximately 10 and 23 % (Table 4). Evidently the ensemble average of the respective timeseries removes uncorrelated “noise” from each record. Analogous results also occur in an analysis of the area-averaged ozone at each level (not shown).

The geographical pattern of the EOFs relates variability between different regions. On each level the variability explained by each EOF is mostly the same sign (Fig. 11) consistent with the relationship between the principal component time series and that of area averaged 30–90° N ozone. At all levels the ozone variability attributed to the first principal component is largest to the north and decreases to the south. The equatorward decrease in the amplitude of the EOF is less at the surface, consistent with the transport of stratospherically derived ozone downwards and southwards along isentropic surfaces. The standard deviation of the ozone variation due to the first EOF reaches 0.6–0.8 ppb at the surface, 1.5–2.0 ppb at 500 hPa and almost 80 ppb.
at 150 hPa. Obviously, at times the amplitude of the first principal component can well exceed these values. At the end of this section we examine the variation of the principal component at selected measurement locations.

Hess and Zbinden (2013) and Zbinden et al. (2006) noted the temporal variability of the ozone record is often similar over widespread geographical regions. The large percent of variability explained by the first EOF, the global nature of this mode, and the fact that it is of the same sign over large regions of the N. H. extratropics demonstrates the connection of the temporal ozone record between geographically distant regions. The vertical coupling between the principal component timeseries of the first EOF and its relation to the 100 hPa ozone flux suggests the root cause of this widespread variability is due to coupled modes of stratosphere–troposphere variability. It seems likely the region west of Ireland where the surface amplitude of the first EOF is comparatively large (Fig. 11) is captured by the measurements at the Mace Head observatory, particularly when the measurements are sampled for "baseline" tropospheric air. This region of large amplitude in the surface EOF pattern (Fig. 11c) helps to explain the relation between the ozone variability sampled at Mace Head and the variability sampled at the high alpine sites over Europe (e.g., Hess and Zbinden, 2013), where the amplitude of the first EOF is comparatively large (not shown). At the European Alpine sites it is appropriate to analyze the simulations at the site elevation and not at the model surface: the site elevation is not resolved in the simulations and they predominantly sample free tropospheric air. The EOF representative of the elevation of these sites does not resemble the surface EOF (Fig. 11c) but more closely resembles the EOF at 500 hPa (Fig. 11b). The 150 hPa and 500 hPa ozonesondes over Canada, Northern Europe and Central Europe also have similar amplitudes of the primary EOF. This suggests the variability between these regions should be highly related as shown in Hess and Zbinden (2013). The amplitude of this mode of variability is less over Japan at both 150 and 500 hPa: the ozone variability over Japan is more likely to be swamped by other modes of variability. As remarked above (also see Hess and Zbinden, 2013) the variability over Japan is not well correlated with the variability in other regions.
At the surface (Fig. 11c) the first EOF is small to occasionally negative over the regions with high emissions: over the Eastern US, Europe and Eastern Asia. In these regions variability is likely governed by local photochemistry and is less influenced by the large-scale processes examined here. In addition, since ozone has a different seasonality between these locations and the more remote regions, annually averaged ozone is likely to reflect different processes. Over the US and Europe surface ozone tends to maximize in summer; in other more remote locations ozone is minimum during the summer months. Over Asia the seasonal variability is likely complicated by the summer monsoon. The EOF pattern as a whole is dominated by variability over remote regions.

The geographic pattern of ozone variability associated with the first EOF reflects known patterns of stratosphere–troposphere exchange (Fig. 11) particularly at the surface. At the surface high regions of variability extend southward over the Eastern Atlantic ocean and the Eastern Pacific Ocean and the Western US. Lin et al. (2012), James (2003) and Sprenger (2003) emphasizes the importance of deep stratospheric ozone intrusions over the Western US coast and Eastern Pacific. Here downward ozone transport from stratospheric sources of ozone can descend along the Eastern flank of the Pacific anticyclone. The outlines of this anti-cyclonic transport of ozone are particularly evident at 1000 and 500 hPa in the first EOF (Fig. 11). Sprenger (2003) also shows that the Atlantic basin is a region of significant stratosphere to troposphere transport of air with a climatological region of deep stratosphere to troposphere exchange extending from southern Greenland to Ireland.

While the surface EOFs calculated for ensemble members are overall qualitatively similar, rather large differences are notable in some locations between the different ensemble members (Fig. 12). In particular, the Western US and Ireland show large differences between different simulations, suggesting the importance of unforced model variability in these regions. For example in one ensemble member ozone variability off the Southwest Coast of the US attributed to the first EOF exceeds 0.6 ppb, while in another member it is less than approximately 0.2 ppb (Fig. 12). The average variabili-
ity for all the ensembles off the Southwest Coast of the US is approximately 0.4 ppb (Fig. 11). The magnitude of the first EOF off Ireland also varies greatly between ensemble members: in one ensemble member the variability attributed to the first EOF is approximately 0.7 ppb near Ireland, in another ensemble member it is close to zero. The ensemble mean is close to 0.2 ppb (Fig. 11).

The time series of the first principal component from each ensemble simulation show little relation during some periods, but during others the ensembles show strong similarities at all levels (Fig. 13). Many of the events where the ensembles show similar behavior appear to be associated with ENSO (Fig. 13): the pronounced negative and positive ozone anomalies during all simulations and all levels during 1966 and 1967 appear to be associated with the negative and positive ENSO indexes that occur 6–12 months earlier; an ozone peak is also common to all levels and all ensembles in 1998–1999 following the El Nino of 1998. The 1998 El Nino event has been linked in the literature to a tropospheric ozone anomaly (Koumoutsaris et al., 2008; Voulgarakis et al., 2011). However, we find the correlation between ENSO and the principal component timeseries is small at all levels (less than 0.23) (not shown). (The correlation is also small on all levels between the ENSO signal and the area averaged 30–90° ozone).

We note that Hsu and Prather (2009) also did not find a relation between ENSO and STE, although Zeng and Pyle (2005) show a strong correlation. Indeed the impact of ENSO on stratospheric circulation statistics via associated changes in stratospheric wave driving, and in particular an increase in the downwards residual velocity at extratropical latitudes during warm ENSO events, provides a mechanism whereby ENSO impacts the stratospheric-tropospheric exchange of ozone (Calvo et al., 2010). However, a careful examination of the ozone perturbations based on a compilation of high-index El Nino events in Calvo et al. (2010) indicates significant ozone perturbations do not persist below about 10 km. In Zeng and Pyle (2005) the ENSO index and STE were only correlated between 1990–2002, a relatively short period compared to the present study. The simulations described in Zeng and Pyle (2005) did not include the forcing due to the QBO or volcanoes included in the present study, forcings that may mask an
underlying ENSO signal. Neu et al. (2014) were not able to isolate the impact of ENSO from that of the Quasi-Biennial Oscillation (QBO) on extratropical tropospheric ozone for the period from 2005 to 2011. The particular period examined in Zeng and Pyle (2005) (1990–2002) was dominated by large events: the 1998–1999 positive ozone anomaly associated with El Nino (Koumoutsaris et al., 2008; Voulgarakis et al., 2011) and the negative anomaly in 2000–2001 that can be associated with La Nina. Figure 12 suggests that only upon occasion is ENSO associated with the strong forcing of ozone anomalies, although, over the long-run the correlation between the two is small.

Sampled at characteristic locations (500 hPa Canadian sites, Mace Head, Lassen and the European alpine sites) (Fig. 14) it is apparent that the ozone variability due to principal component timeseries explains a substantial fraction of the overall simulated variability. Over the course of the simulation the correlation between ozone variability and the principal component timeseries is: 0.97 for all ensemble simulations over the Canadian 500 hPa ozonesonde sites; between 0.64–0.80 for the ensemble simulations over the European alpine sites; between 0.48 and 0.70 for the ensemble simulations at Mace Head and between 0.09 and 0.74 for the ensemble simulations at Lassen. Note however at Lassen three of the simulations have correlations above 0.60. To better sample baseline ozone conditions at Mace Head we have sampled the model at the point 10° W and 5° N of the actual observatory location in Fig. 14. The correlation between the model and measurements has been described above (Sect. 3; also see Table 2)

Note in particular, the ozone increase between the measured ozone minimum in the early 1990s and the ozone maximum near 1998–1999 at the disparate locations shown in Fig. 14 is not only captured in the simulated ozone record but also in the principal component timeseries. At the four sites in Fig. 14, the ozone jump during the 1990s (defined here as the maximum minus minimum annually averaged ozone during the 1990s, where the ozone has not been detrended) is 6.6 ppb at Mace Head, 7.4 ppb at the European Alpine sites, 9.1 ppb at Lassen and 12.6 ppb at the Canadian ozonesonde stations. The simulations capture approximately 50% of the measured in-
crease at all stations, ranging from 47% over the Canadian stations to 56% at Mace Head. A regression against methane over the entire model simulation show that increases in methane explain a relatively small fraction of the ozone increase during the 1990s at these sites: 0.8 ppb decade$^{-1}$ at the Lassen site, 0.27 ppb decade$^{-1}$ over Canada, 0.39 ppb decade$^{-1}$ at Mace Head and 0.27 ppb decade$^{-1}$ at the European Alpine sites. In fact, much of the simulated ozone increase at these sites can be traced to the principal component time series at these sites. Over Canada, the European Alpine sites, Lassen and Mace Head changes in the principal component account for 100, 68, 49 and 43% of the simulated ozone jump, showing that much of the jump during this period can be traced to changes in this global mode of variability. As argued above, the vertical correlation of the ozone principal component timeseries from the surface through the lower stratosphere, as well as their correlation with the 100 hPa ozone flux suggests changes in the principal component are consistent with changes in the STE of ozone. Hess and Zbinden (2013) show that much of the measured ozone change during the 1990s at a variety of tropospheric sites could be traced to increases in the stratospheric portion of ozone.

Ozone increases measured on the west coast of the US at a variety of stations have been ascribed to increasing Asian emissions (e.g., Cooper et al., 2010; Parrish et al., 2012). However, the results here show substantial ozone increases have occurred during the 1990s over a wide variety of sites, many of them substantially removed from Asian emissions. In addition, the simulations, with no change in emissions, capture almost 50% of the observed ozone jump during the 1990s, including the changes at Lassen. This suggests that a large portion of the measured jump is not due to changes in emissions, but can be traced to changes in a global mode of ozone variability. This emphasizes the difficulty in the attribution of ozone changes, but also the importance of understanding the importance of natural variability in isolating the role of emissions in modifying ozone concentrations.
5 Conclusions

We have analyzed an ensemble of four free running simulations from 1953–2005 using the Whole Atmosphere Community Climate Model (WACCM). The simulations are forced by time changes in observed sea-surface temperatures, concentrations of greenhouse gases (including methane), stratospheric ozone depleting species, an externally forced quasi-biennial oscillation, solar variability and stratospheric sulfate surface area density (SAD). In the stratosphere WACCM employs a sophisticated chemical mechanism. In the troposphere only the basic tropospheric NOx-CH4 chemistry is used, where the NOx surface emissions remain interannually constant. The relation between tropospheric interannual ozone variability and lower stratospheric interannual ozone variability is analyzed from 30–90° N.

Despite the simplicity of the tropospheric chemistry, the simulations capture the measured N. H. background ozone interannual variability to a surprising extent. Particularly for the period after 1990 the simulated detrended ozone is significantly correlated with the 500 hPa ozonesonde measurements over Northern Europe, Central Europe and Canada (but not over Japan) and over surface measurement sites at Mace Head and the alpine sites over Europe (but not at Lassen or Arkona). We argue that the Arkona site (in Germany) is influenced by fresh anthropogenic emissions that may not be accurately captured with the simple tropospheric chemistry in these simulations. While the simulation appears to capture some aspects of the Lassen record, the large ozone trend in the measurements makes a simple comparison difficult. Prior to 1990 there are fewer measurements and the reliability of the measurements decreases. However, even during this period the simulated ozone record is significantly correlated with the measured record in a number of locations.

It is not a foregone conclusion that the simulated detrended ozone should be correlated with the measurements in the first place, as the model dynamics are internally calculated. The fact that the model and observations are significantly correlated implies the importance of external forcing in determining the ozone variability. Indeed, the
interannual simulated ozone record shows periods when the ozone variability appears to have little relation between the different model ensemble simulations. At these times the applied external forcing does not exert a strong control on the ozone variability and the variability is likely dominated by internal model variability. However, the periods when the ensemble simulations show little similarity are punctuated by times when the variability between all the ensembles show strong agreement suggesting the importance of a common forcing mechanism. At the sites examined the correlation between the different detrended ensemble members ranges from approximately 0.24–0.44 at the surface to 0.4–0.5 at 500 and 150 hPa. The decade of the 1990s, in fact, may be particularly impacted by external forcing due to the influence of Mt Pinatubo during the early part of the decade and the impact of the 1998 El Nino toward the end of the decade.

There appears to be some association between the punctuated periods when all ensembles show strong agreement and ENSO. However, we do not find that these periods occur with all ENSOs, even if the ENSO is particularly strong. We find little relation between the ENSO index and large-scale tropospheric ozone variability over the long-term record. We argue the length of the simulated record in the current study and the inclusion of volcanic and QBO forcing may explain the difference between this study and earlier work (e.g., Zeng and Pyle, 2005).

The simulated curves of tropospheric ozone vs. methane at a number of sites show a relatively rapid ozone increase prior to 1970, a subsequent slow down in the rate of ozone increase from 1970–1985, but subsequent increased ozone growth after approximately 1990. The measured curves are strongly impacted by changes in ozone precursor emissions and thus despite some similarities with the simulations remain difficult to interpret with respect to STE. The ensemble average tropospheric ozone record can largely be explained as a linear combination of the 30–90° area averaged 100 hPa ozone flux and the global methane concentration. We use the former quantity as a proxy for STE. The long-term non-linear rate of ozone increase with respect to methane can be explained by changes in the downward ozone flux across the 100 hPa
level. As expected from the imposed change in greenhouse gas forcing, the strength of the residual circulation increases throughout the simulations. This alone would act to increase the downward extratropical N. H. stratospheric ozone flux with a resulting increase in tropospheric ozone; however, stratospheric ozone depletion counteracts this. As a consequence the 100 hPa ozone flux decreases between approximately 1970–1990 and the rate of growth of tropospheric ozone with respect to methane slows. Subsequent to Mt Pinatubo ozone increases in the extratropical N. H. stratosphere. This acts to increase the 100 hPa stratospheric ozone flux with resulting increases in tropospheric ozone. The multivariate linear regression of ozone against methane and the 100 hPa ozone flux gives an approximate 20 % sensitivity to changes in methane; the average sensitivity to the ozone flux is 19 % at 500 hPa and 11 % at the surface. Without ozone depletion the approximate 15 % increase in the simulated residual circulation from 1952–2005 would have resulted in a consequent 1.5 % increase in surface ozone in the N. H. extratropics. Ozone depletion has reduced these increases by approximately half. Extrapolating these changes, a 30 % increase in the ozone flux by 2100 (Hegglin and Shepherd, 2009) would result in 3 % increase in surface ozone and a 6 % increase in 500 hPa ozone. In contrast the satellite based study of Neu et al. (2014) suggests an increase of the stratospheric circulation by 40 % would lead to a tropospheric ozone increase of 2 %.

On an interannual timescale changes in the ensemble averaged 100 hPa ozone flux (averaged from 30–90° N) explains 70 % of the ensemble averaged extratropical detrended ozone variability at 500 hPa and 55 % of the ensemble averaged detrended ozone variability at the surface. In regions of large emissions (e.g., Arkona) the variability explained is much less. Sampling “baseline” air just to the northwest of Mace Head suggests variations in the ozone flux explain almost 25 % of the variability of the “baseline” tropospheric ozone variability at Mace Head. The first empirical orthogonal function (EOF) of 30–90° N ozone variability explains from 40 % of the ozone variability at the surface to over 80 % of the ozone variability at 150 hPa. The spatial pattern of this EOF at the surface is consistent with expected pat-
terns of stratosphere–troposphere exchange, with a maximum off the west coast of the US. The EOF also shows relatively high amplitudes off the west coast of Ireland. The EOF has a small amplitude over regions with significant emissions of ozone precursors. The principal component timeseries associated with the first EOF are highly correlated in the vertical. The ensemble average principal component time series is also highly correlated with the ensemble averaged \(100\) hPa ozone flux and is highly correlated with the ensemble averaged \(30–90^\circ\) ozone average.

The interannual ozone variability at a number of individual sites over the course of the model simulation is largely explained by the variability in the principal component of the global EOF. At the stations examined in detail (the Canadian ozonesonde stations, the European Alpine sites, Lassen and Mace Head) the simulated ozone increase during the 1990s is about 50% of the measured increase. Over Canada, the European Alpine sites, Lassen and Mace Head changes in the principal component account for 100%, 68%, 49% and 43% of the simulated ozone jump, showing that much of the jump during this period can be traced to changes in a global mode of ozone variability. This suggests that a large portion of the ozone increase in the 1990s as measured at a number of sites is not due to changes in emissions, but can be traced to changes in a global mode of ozone variability. This emphasizes the difficulty in the attribution of ozone changes, and the importance of natural variability in understanding the trends and variability of ozone (see Lin et al., 2014). The mode of variability analyzed here shows strong stratosphere–troposphere coupling, demonstrating the importance of the stratosphere in the attribution of tropospheric ozone variability.

Despite the simplicity of the tropospheric chemistry used in these simulations, the simulations match the observed tropospheric variability to a large extent over locations sampling background tropospheric air. It is expected that the introduction of additional hydrocarbon chemistry as well as episodic emission variability (e.g., biomass burning) will introduce modes of variability not captured here as well as possibly dampen the basic modes of ozone variability analyzed above. Future simulations are necessary to fully explicate the importance of episodic emission variability and of the variability
associated with hydrocarbon chemistry including that of biogenic emissions. However, based on the agreement between these simulations and measurements we hypothesize that the base state modes of variability isolated here are fundamental to the coupled troposphere-stratosphere chemical system. The results obtained here are largely consistent with those in Hess and Zbinden (2013), where a sophisticated tropospheric mechanism is employed along with a methodology for tagging stratospheric ozone. Hess and Zbinden (2013) also found the exchange of ozone from the stratosphere to the troposphere explains a large fraction of extratropical ozone variability, even at the surface, and this variability operates on hemispheric spatial scales.

The high variability explained by this global mode and the fact that it is of the same sign over large regions of the N. H. extratropics demonstrates the relationship of the temporal ozone record between geographically distant tropospheric regions. Hess and Zbinden (2013) and Zbinden et al. (2006) noted the temporal variability of the ozone record is often similar over widespread geographical regions. The vertical coupling between the principal component timeseries of the first EOF and its relation to the 100 hPa ozone flux suggests the root cause of this widespread variability is due to coupled modes of stratosphere–troposphere variability.

It is perhaps surprising that the stratosphere should be so important in explaining interannual ozone variability. While early work suggested that much of the tropospheric ozone distribution can be explained with a stratospheric ozone source and a surface sink (Levy et al., 1985), for the past several decades, a newer paradigm emerged which de-emphasized the role of the stratosphere. It was recognized that tropospheric photochemical ozone production is almost an order of magnitude larger than the input from the stratosphere e.g., (Stevenson et al., 2006). However, the stratospheric source of ozone is widespread (Liang et al., 2009) and the lifetime of ozone is relatively large in the upper and middle stratosphere, allowing ozone from the stratosphere to be transported throughout the troposphere (Liang et al., 2009). By contrast, ozone production can be large near the surface, but on average this is compensated for by photochemical and surface ozone loss. The stratospheric source is particularly important in explain-
ing the interannual variability of ozone away from regions of immediate photochemical production.

These results suggest the difficulty in the attribution of ozone changes without understanding the root causes of the natural variability of ozone. Simply examining changes in ozone precursor emissions, even on the decadal timescale, is insufficient to link changes in ozone to changes in emissions. A full attribution of ozone variability may require more sophisticated models with a good resolution of stratospheric processes.

Acknowledgements. P.G. Hess would like to acknowledge NSF grant #1042787 for supporting this work. Work at LLNL was performed under the auspices of the US Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344 and supported by the Atmospheric Radiation Measurement Program of the Office of Science at the US Department of Energy. The CESM project is supported by the National Science foundation and the US Department of Energy. The National Center for Atmospheric Research is operated by the University Corporation for Atmospheric Research under sponsor ship of the National Science Foundation. We acknowledge the World Ozone and Ultraviolet Radiation Data Centre (WUDC) for providing the ozonesonde data, the US National Park Service for providing Lassen NP ozone data, A. Volz-Thomas for the Arkona data, and to D. Parrish for providing access to his surface composited ozone datasets.

References


Table 1: Measurement sites used in this paper.

<table>
<thead>
<tr>
<th>Measurement sites</th>
<th>Platform</th>
<th>Lon.</th>
<th>Lat.</th>
<th>Elev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>Ozonesonde</td>
<td>62°W</td>
<td>82°N</td>
<td>NA</td>
</tr>
<tr>
<td>Alert</td>
<td>Ozonesonde</td>
<td>94°W</td>
<td>59°N</td>
<td>NA</td>
</tr>
<tr>
<td>Churchill</td>
<td>Ozonesonde</td>
<td>85°W</td>
<td>60°N</td>
<td>NA</td>
</tr>
<tr>
<td>Eureka</td>
<td>Ozonesonde</td>
<td>60°W</td>
<td>53°N</td>
<td>NA</td>
</tr>
<tr>
<td>Goose Bay</td>
<td>Ozonesonde</td>
<td>95°W</td>
<td>75°N</td>
<td>NA</td>
</tr>
<tr>
<td>Resolute</td>
<td>Ozonesonde</td>
<td>5°E</td>
<td>52°N</td>
<td>NA</td>
</tr>
<tr>
<td>Central Europe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deblit</td>
<td>Ozonesonde</td>
<td>11°E</td>
<td>48°N</td>
<td>NA</td>
</tr>
<tr>
<td>Hoherpeissenberg</td>
<td>Ozonesonde</td>
<td>21°E</td>
<td>52°N</td>
<td>NA</td>
</tr>
<tr>
<td>Legnino</td>
<td>Ozonesonde</td>
<td>14°E</td>
<td>52°N</td>
<td>NA</td>
</tr>
<tr>
<td>Lindenberg</td>
<td>Ozonesonde</td>
<td>8°E</td>
<td>47°N</td>
<td>NA</td>
</tr>
<tr>
<td>Uccle</td>
<td>Ozonesonde</td>
<td>4°E</td>
<td>51°N</td>
<td>NA</td>
</tr>
<tr>
<td>Arkona</td>
<td>Surface</td>
<td>13°E</td>
<td>54°N</td>
<td>42 m</td>
</tr>
<tr>
<td>Jungfraujoch</td>
<td>Surface</td>
<td>8°E</td>
<td>47°N</td>
<td>3680 m</td>
</tr>
<tr>
<td>Mace Head</td>
<td>Surface</td>
<td>10°W</td>
<td>53°N</td>
<td>25 m</td>
</tr>
<tr>
<td>Zugspitze</td>
<td>Surface</td>
<td>11°E</td>
<td>47°N</td>
<td>2860 m</td>
</tr>
<tr>
<td>Japan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kagoshima</td>
<td>Ozonesonde</td>
<td>131°E</td>
<td>32°N</td>
<td>NA</td>
</tr>
<tr>
<td>Sapporo</td>
<td>Ozonesonde</td>
<td>141°E</td>
<td>43°N</td>
<td>NA</td>
</tr>
<tr>
<td>Tateno</td>
<td>Ozonesonde</td>
<td>140°E</td>
<td>36°N</td>
<td>NA</td>
</tr>
<tr>
<td>Northern Europe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ny Alesund</td>
<td>Ozonesonde</td>
<td>12°E</td>
<td>79°N</td>
<td>NA</td>
</tr>
<tr>
<td>Scoresbystrand</td>
<td>Ozonesonde</td>
<td>22°W</td>
<td>70°N</td>
<td>NA</td>
</tr>
<tr>
<td>Sodankyla</td>
<td>Ozonesonde</td>
<td>26°W</td>
<td>67°N</td>
<td>NA</td>
</tr>
<tr>
<td>United States</td>
<td>Surface</td>
<td>122°W</td>
<td>41°N</td>
<td>1769 m</td>
</tr>
</tbody>
</table>

1 Canadian ozonesonde sites are averaged together.
2 Ozonesonde data is from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC).
3 Central European (C. Europe) ozonesonde sites are averaged together.
4 Arkona data courtesy of Andreas Voit-Thomas.
5 For the Jungfraujoch and Zugspitze measurements please see Gilio et al. (2010).
6 Jungfraujoch and Zugspitze (JF/ZU) measurements are averaged together.
7 For the Mace Head measurements please see Donev et al. (2013).
8 Japanese ozonesonde sites are averaged together.
9 Northern European (N. Europe) ozonesonde sites are averaged together.
10 Provided by the National Park Service.
Table 2. Comparison between ensemble mean simulated and measured ozone at various sites (see Table 1).

<table>
<thead>
<tr>
<th>Stations</th>
<th>Years</th>
<th>Meas mean (ppbv)</th>
<th>Model bias (ppbv)</th>
<th>Correlation&lt;sub&gt;3&lt;/sub&gt;</th>
<th>Ensemble correlation&lt;sup&gt;4&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;1990</td>
<td>&gt;1990</td>
</tr>
<tr>
<td>150 hPa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>15 May 1966–15 Sep 2005</td>
<td>627.1</td>
<td>13.6</td>
<td>0.46</td>
<td>0.54</td>
</tr>
<tr>
<td>N. Europe</td>
<td>15 Dec 1988–15 Sep 2005</td>
<td>602.9</td>
<td>6.9</td>
<td>–</td>
<td>0.57</td>
</tr>
<tr>
<td>C. Europe</td>
<td>15 Feb 1967–15 Sep 2005</td>
<td>423.9</td>
<td>11.2</td>
<td>0.64</td>
<td>0.32</td>
</tr>
<tr>
<td>Japan</td>
<td>15 May 1969–15 Sep 2005</td>
<td>242.4</td>
<td>–1.1</td>
<td>0.12</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500 hPa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>15 May 1966–15 Sep 2005</td>
<td>54.6</td>
<td>–2.1</td>
<td>0.40</td>
<td>0.57</td>
</tr>
<tr>
<td>N. Europe</td>
<td>15 Dec 1988–15 Sep 2005</td>
<td>57.9</td>
<td>–5.9</td>
<td>–</td>
<td>0.73</td>
</tr>
<tr>
<td>C. Europe</td>
<td>15 Feb 1967–15 Sep 2005</td>
<td>60.0</td>
<td>–6.1</td>
<td>0.16</td>
<td>0.66</td>
</tr>
<tr>
<td>Japan</td>
<td>15 May 1969–15 Sep 2005</td>
<td>58.0</td>
<td>–4.2</td>
<td>–0.07</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JF/JZUG</td>
<td>15 Oct 1978–15 Sep 2005</td>
<td>50.9</td>
<td>–2.4</td>
<td>–0.16</td>
<td>0.66</td>
</tr>
<tr>
<td>Lassen</td>
<td>15 Sep 1988–15 Sep 2005</td>
<td>40.4</td>
<td>5.2</td>
<td>–</td>
<td>0.25</td>
</tr>
<tr>
<td>Mace Head</td>
<td>15 Jan 1988–15 Sep 2005</td>
<td>38.5</td>
<td>–6.1</td>
<td>–</td>
<td>0.65</td>
</tr>
<tr>
<td>Arkona</td>
<td>15 May 1957–15 Nov 2002</td>
<td>28.7</td>
<td>–3.0</td>
<td>0.63</td>
<td>0.01</td>
</tr>
</tbody>
</table>

1 Years over which measurements and the simulation are evaluated.
2 The bias is evaluated between 1990 and 2005.
3 Correlation between 12-month smoothed detrended ensemble mean ozone and 12-month smoothed detrended measurements before and after 1990. Significant correlations (at 99%) are in bold.
4 Correlation between ensemble members: lowest correlation, median correlation, high correlation. Median correlations significant at 95% in bold. Correlations are between 12-month smoothed records.
5 Jungfraujoch/Zugspitze.
Table 3. Sensitivity coefficients (percent change in ozone to a percent change variable) between 12 month smoothed normalized simulated ozone and normalized globally averaged methane and the normalized lagged 30–90° N ozone flux. Coefficients given for the lag with the smallest chi squared. Variables are normalized by their averaged value from 1980–1985.

<table>
<thead>
<tr>
<th>Stations</th>
<th>Sensitivity CH₄ Flux</th>
<th>Lag¹ (months)</th>
<th>Corr²</th>
<th>CHISQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 hPa Canada</td>
<td>0.15</td>
<td>0.25</td>
<td>5</td>
<td>0.86</td>
</tr>
<tr>
<td>N. Europe</td>
<td>0.16</td>
<td>0.24</td>
<td>6</td>
<td>0.83</td>
</tr>
<tr>
<td>C. Europe</td>
<td>0.18</td>
<td>0.17</td>
<td>5</td>
<td>0.77</td>
</tr>
<tr>
<td>Japan</td>
<td>0.22</td>
<td>0.08</td>
<td>5</td>
<td>0.47</td>
</tr>
<tr>
<td>30–90° N Aver⁵</td>
<td>0.17</td>
<td>0.19</td>
<td>4</td>
<td>0.84</td>
</tr>
<tr>
<td>Surface JF/JUG</td>
<td>0.20</td>
<td>0.13</td>
<td>6</td>
<td>0.67</td>
</tr>
<tr>
<td>Lassen</td>
<td>0.15</td>
<td>0.14</td>
<td>5</td>
<td>0.67</td>
</tr>
<tr>
<td>Arkona</td>
<td>0.45</td>
<td>0.04</td>
<td>5</td>
<td>0.11</td>
</tr>
<tr>
<td>Mace Head</td>
<td>0.20</td>
<td>0.09</td>
<td>5</td>
<td>0.28</td>
</tr>
<tr>
<td>30–90° N Aver⁵</td>
<td>0.21</td>
<td>0.11</td>
<td>6</td>
<td>0.73</td>
</tr>
</tbody>
</table>

¹ Lag in months between the ozone record and the 30–90° N averaged ozone flux resulting in the smallest regressed chi-squared. The lag is measured as the number of months by which the ozone concentration lags the ozone flux.
² Correlation is between the regressed ozone record and the simulated ozone record after removing the regressed dependence on methane from each (see text). Values significant at the 90% level are shown in bold.
³ There is no well defined minimum chi-squared at Arkona. We give coefficients at five months.
⁴ Correlation is 0.49 for the point 10° W and 5° N of Mace Head.
⁵ Average from 30–90° N.
Table 4. Explained variances and correlations between EOFs on various levels.

<table>
<thead>
<tr>
<th>Level</th>
<th>Explained variance$^1$</th>
<th>Correlation 150 hPa PC$^2$</th>
<th>Correlation O$_3$ flux 100 hPa$^3$</th>
<th>Lag$^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>150 hPa</td>
<td>79–85 %</td>
<td>NA</td>
<td>0.85 (0.43)</td>
<td>2(2)</td>
</tr>
<tr>
<td>500 hPa</td>
<td>71–77 %</td>
<td>0.78</td>
<td>0.85 (0.50)</td>
<td>5(5)</td>
</tr>
<tr>
<td>1000 hPa</td>
<td>40–48 %</td>
<td>0.66</td>
<td>0.73 (0.34)</td>
<td>6(6)</td>
</tr>
</tbody>
</table>

$^1$ Range of variances explained by the 1st EOF over the model ensembles.

$^2$ Temporal correlation between principal components at 1000 hPa and 500 hPa and the principal component at 150 hPa. Correlation is computed between levels of the same ensemble simulation; however, the overall correlation coefficient comprises the relationship for all ensembles. All correlations are significant at the 99 % level.

$^3$ Temporal lagged correlation between principal components on various pressure levels and the 30–90° N averaged ozone flux at 100 hPa. The correlation without parenthesis is between the the ensemble averaged principal component and the ensemble averaged 30–90° N 100 hPa ozone flux. The correlation in parenthesis is between the principal component and the 30–90° N 100 hPa ozone flux of each individual ensemble member. All correlations are significant at the 99 % level.

$^4$ Lag (months) of the maximum correlation between the ozone flux and the principal component: without parenthesis for the ensemble average; with parenthesis for individual ensemble members. The lag is measured as the number of months by which the ozone concentration lags the ozone flux.
Figure 1. Model Forcings. Concentrations of CH$_4$ (green, left axis) and CFC-11 (blue, right axis) globally area-averaged at lower boundary and used to force the WACCM simulations.
Figure 2. Ozone deviations (ppb) averaged from 30–90° N for each WACCM ensemble member (colored) and the deviation averaged over all ensemble members (black) at: (a) 150 hPa, (b) 500 hPa and (c) surface. Monthly ozone deviations are smoothed over 12 months. Deviations are from ozone averaged 1 January 1990–31 December 1994.
Figure 3. Simulated and measured ozone deviations (ppb) at 150 hPa averaged over the (a) Canadian ozonesonde sites, (b) the Central European ozonesonde sites, (c) the Japanese ozonesonde sites and (d) the Northern European ozonesonde sites. The simulated ensemble average is given as the bold black line, the thin black lines bracket the maximum and minimum ensemble ozone deviation, the measured average is given as the red line, the blue dots give the measured ozone deviation for each site comprising the regional average. Colored bars indicate when each measurement site (color coded as indicated on right) made sufficient measurements to calculate an annual ozone concentration: solid lines indicate an ECC measurement and dotted lines a BrewerMast ozonesonde measurement. The black and red lines at the bot- tom give the simulated (black) and measured (red) standard deviation of ozone (ppb) calculated across all sites within each region. Numbers in the upper right give the model-measurement correlation of the average ozone within each region prior to 1990 (left) and after 1990 (right). Correlations use detrended data. Significant correlations at the 95 % level are starred. Monthly ozone deviations are smoothed over 12 months. Deviations are from ozone averaged 1 January 1990–31 December 1994. The simulated ozone record uses monthly averaged output at each measurement site throughout the simulation with no attempt to duplicate the actual temporal sampling of the measurements.
Figure 4. As in Fig. 3, but at 500 hPa averaged over the (a) the Canadian ozonesonde sites, (b) the Central 500 hPa European ozonesonde sites, (c) the Japanese ozonesonde sites and (d) the Northern European ozonesonde sites.
Figure 5. As in Fig. 3 but for the surface simulated and measured ozone deviations (ppb): (a) averaged for the Jungfraujoch and Zugspitze sites; (b) the Lassen site. The bottom bars in (a) indicate the years for which an annually averaged measurement was available at the Jungfraujoch and Zugspitze sites.
Figure 6. As in Fig. 3, but for (a) the surface measurements at Arkona and (b) the surface measurements at Mace Head.
Figure 7. The (a) vertical residual velocity ($w^*$, m s$^{-1}$) and the (b) ozone flux (Tg yr$^{-1}$) averaged on the 100 hPa surface between 30 and 90° N for each ensemble simulation (colored). The ensemble average fields are fit cubically and shown in black. A 12 month smoothing is used for all fields.
Figure 8. (a) Simulated and (b) measured cubic fits of normalized ozone vs. global surface methane concentration for various long-term tropospheric measurement sites: the regional average of the Canadian (Can.), Central European (Ceur.) and Japanese (Jap.) ozonesonde sites at 500 hPa, the average of the Jungfraujoch and Zugspitze sites (Alpine S.) and the Arkona surface site. Ozone is normalized by its 1980–1985 concentration at each site. Globally averaged methane is from the WACCM simulation. The year corresponding to the methane concentration is given. Simulated ozone is the ensemble mean.
Figure 9. Simulated normalized ensemble mean ozone against global methane concentrations (blue squares), the cubic fit of normalized ozone against methane (red line) and the regressed fit (black line). Ozone is averaged from 30–90° N at (a) the surface, (b) 500 hPa. The year corresponding to the methane concentrations is also shown.
Figure 10. Ozone deviations (ppb) averaged from 30–90° N for each of the four WACCM ensemble members at 150 hPa (black), 500 hPa (blue) and the surface hPa (green). The linear dependence on global methane has been removed from the ozone records at 500 and 1000 hPa. Monthly ozone deviations are smoothed over 12 months. Deviations are from ozone averaged over the entire simulation. Note the different scales for each level.
Figure 11. Normalized first EOF component of detrended ozone at (a) 150 hPa, (b) 500 hPa and (c) surface. Shown is the average for all four ensembles of the EOF multiplied by the standard deviation of the principal component. The absolute value of the result shows the variability of ozone (ppb) expected due to variations in the first EOF component, the sign of the result shows the relation between variability in different locations.
Figure 12. As in Figure 11, but for the normalized first EOF component of detrended ozone at the surface for two different ensemble members.
Figure 13. Timeseries of the principal component for the first EOF of ozone from 30–90° N. for each ensemble simulation (color) and for the ensemble mean (black) at (a) 150 hPa, (b) 500 hPa and (c) surface. (d) The ENSO index is shown in the lower panel.
Figure 14. 12 month smoothed ozone deviations (ppb) for (a) the 500 hPa Canadian ozonesondes, and the surface sites at (b) Mace Head, (c) Lassen, and (d) the European Alps (note different scales in each figure): detrended measurements (red), ensemble average detrended ozone (black), the time variation of the EOF (blue), where the vertical blue lines bracket the range of the EOF over the ensemble members and the blue dot gives the ensemble average EOF. In each case ozone deviations are detrended against globally averaged methane over the common range of simulated and measured ozone.