ATP Ms. 2013-437:

"Decadal-scale responses in middle and upper stratospheric ozone from SAGE II version 7 data" by E. E. Remsberg

General Comments:

This is an interesting but unconvincing statistical analysis and interpretation of decadal variations and long-term trends in SAGE II ozone profile data covering 1984-2005 (as well as HALOE ozone data from 1991 to 2005). With respect to scientific significance, the analysis of version 7 SAGE II data is substantially new but the concepts, ideas, and methods are not (see comment 1 below). More importantly, a major conclusion (that the SAGE II and HALOE results for the solar cycle-like response during the 1991-05 period are in agreement with each other and with representative 2D model estimates, "after taking into account the geophysical units used in the simulation studies"; p. 17, lines 6-9) is not really supported by the evidence presented in the manuscript (comment 2 below). Moreover, the SAGE solar cycle-like response during the 1984-98 period and that derived from alternate data sets (SBUV) lead to a very different conclusion, which is not mentioned in the abstract or Conclusions section. The significance is therefore rated as fair (3). With respect to scientific quality, the statistical approach that is applied and the short lengths of the time series that are analyzed to estimate the solar cycle-like response (15 years) lead to complications, especially due to interfering effects of the Pinatubo aerosol injection event and interannual dynamical variability (comment 1 below). The results are also not discussed in a balanced way in the final section in my opinion, especially with respect to the main conclusion noted above. The quality is therefore also rated as fair (3). The presentation quality is generally good(2) although some of the 15 figures could possibly be eliminated.

Specific Comments:

1. Although the analysis of version 7 SAGE II data is new, like the previous study of Remsberg and Lingenfelser (2010), the statistical approach that is applied is unconventional: Rather than applying a multiple linear regression (MLR) model containing physical explanatory variables (QBO, solar cycle, volcanic aerosol, ENSO, EESC or piecewise linear trends), the MLR model contains only general periodic and linear trend terms. In addition, the available data record (1984-2005) is divided into two shorter overlapping 15-year segments (1984-98; 1991-2005) and results are presented for the separate segments rather than for the whole 22-year record. This approach has a number of negative consequences. First, the anomalously large solar cycle-like responses in the middle stratosphere at northern low latitudes during the 1991-05 period (Figures 2 and 10) are an artificial consequence of "end point anomalies" due to the Pinatubo aerosol injection at the beginning of that segment, which is not accounted for in the MLR model (discussion at top of p. 10). Also, the apparent large positive response in the equatorial middle stratosphere during the 1984-98 period occurs at a lag of several years (Figures 5 and 6), possibly because of changes in the amplitude of the QBO that are not accounted for in the MLR model (discussion in lines 10-15 on p. 10). This is repaired in Figure 9 by adjusting the amplitudes of Figure 5 to ensure a constant phase with respect to the observed solar cycle. However,

the problem would have been naturally avoided if an MLR model with a solar cycle term had been applied. Finally, because of the short record lengths, the results are sensitive to interannual dynamical variability. As discussed beginning at the bottom of p. 11, a major difference between the final SC-like responses for the two segments occurs near the tropical stratopause where the response for 84-98 is nearly 4% (Figure 9) while that for 91-05 shows a distinct minimum of about 2% (Figure 10). As shown in Figure 11, the reduced amplitude during 91-05 is likely because of dynamically induced ozone minima occurring in association with sporadic stratospheric sudden warming episodes that occurred for most years during the 80's but were nearly absent during 93-98, thereby tending to cancel the solar signal at this location for the 91-05 segment. If the full record had been analyzed, this problem would have been mitigated.

2. A major conclusion given in lines 6-9 on p. 15 of the Conclusions section states: "The SC-like response profiles from the SAGE II and HALOE ozone of the low latitudes" during the 1992-2005 period "agree well with representative model profiles, after taking into account the geophysical units used in the simulation studies." This statement is not really supported by the detailed analysis presented in section 3, especially evidence that the reduced SAGE II response in the tropical upper stratosphere during this period (which is a factor of 2 less than obtained during the 84-98 period) is likely due to interannual dynamical variability during the short measurement record rather than to solar forcing (Figure 11 and associated discussion beginning on p. 11, line 22). If this reduction is due to dynamical variability, then the tendency to agree with the shape of the 2D model shown in Figure 13, is only fortuitous. Also, the 92-05 period is selected to yield the best possible agreement. When the 84-98 and 91-05 periods are used, there is no agreement as shown in Figure 12. The author notes that this is the case on p. 13: "Still, the SAGE II response is smaller in the upper stratosphere for 1992-2005 than for 1984-1998 (Fig. 12), perhaps related to decadal-scale differences for the net circulations mentioned earlier." The correct conclusion given in section 5 should therefore have been that there is still a problem in the upper stratosphere between the SAGE II ozone SC response and that which is estimated in most 2D and 3D models. This is further verified when alternate data sets (SBUV) are considered (WMO, 2007). Some discussion of SBUV results should have been given in the paper with appropriate references.

3. Along the same lines, the summary given in the abstract is a very optimistic view of reality: "The shapes of their (SAGE II and HALOE) SC-like response profiles agree well for a time series from late 1992-2005, or after excluding the first 14 months of data following the Pinatubo eruption. Max minus min, SC-like responses from the SAGE II and HALOE time series vary from 2 to 4% and from 0 to 2%, respectively, and their differences in the upper stratosphere can be accounted for using the analyzed, SC-like response of the HALOE temperatures." It is not mentioned here that the reduced response in the upper stratosphere during this short period is likely due to interannual dynamical variability. Also, according to the comparison in Figure 13, the SAGE II response is still 1 to 2% larger than the HALOE response at most altitudes. This difference can't easily be explained by adjusting the SAGE measurements to pressure levels using observed temperature variations because the annual mean solar cycle change in tropical temperature is only statistically

significant in the uppermost stratosphere (above 45 km) and in the lowermost stratosphere (below 25 km) (see Fig. 1c of Frame and Gray, 2010). The zone between 25 and 45 km still shows differences of up to 2% in Figure 13, which are not easy to explain by such adjustments. The paper does not show that this is the case. There are still, therefore, substantial unexplained differences between the HALOE and SAGE II SC results even when a reduced (92-05) time period is considered and a blind eye is turned toward the 84-98 SAGE II results and the SBUV results.

Technical Corrections:

4. Even though this analysis indicates that the linear trend differs between the first and the last part of the SAGE II record, it would have been possible to analyze the entire 84-05 period to estimate the overall SC-like ozone response if a piecewise linear trend term had been included in the regression model. This is a method that makes no assumptions about the origins of the trends.

5. P. 3, line 2: Frame and Gray (2010) did not analyze ozone data, only reanalysis meteorological data, I believe.

6. The reason for limiting the SAGE II analysis to altitudes of 27.5 km and higher (Figure 2 and following figures) is not discussed. It would have been useful to analyze data down to 20 km to allow more direct comparisons with alternate analyses (e.g., Randel and Wu 2007), which have found evidence for a lower stratospheric response. However, it would be necessary to carefully consider aerosol contamination effects when considering data to altitudes as low as 20 km.

7. Figures 12-15. Two standard deviation error bars should have been shown on these figures, at least for representative pressure levels.