We thank the anonymous reviewer for their helpful comments on our paper. In the following, the original reviewer comment is repeated in italics, followed by our response.

_The main purpose of paper is the measurement intercomparison between the different ground-based instruments so that these measurements can be used as a reliable source for the validation of the ACE satellite data. However, given that the results from the comparison between the ground-based and satellite data sets is quite prominently presented in the abstract, the actual validation part seems somewhat hidden in text (see also specific comments). Why? I would have rather expected a separate section on the validation results, rather then finding it spread through the text._
E.g. Section 6.2, the comparison between ground-based and satellite data sets is also described here under each of the Sections for each year. Might make sense to separate that part out or at least to reflect it somewhat clearer in the title that the validation of the ACE data sets is discussed here.

We have reordered Section 6 to put more emphasis on the satellite comparisons. We have grouped the ozone and NO2 ground-based comparisons together in Section 6.2 (now titled "Comparisons between ground-based instruments") and the ozone and NO2 satellite comparisons together in Section 6.3 (now titled "Comparisons with satellite instruments"). This has also resulted in the reordering of Figures 15-20.

Specific comments and suggestions:

We have made the typographical and clarity corrections found by the reviewer. If the changes require comment they are included below.

Page 16286, line 11: To be consistent, add (MAESTRO) after "onboard ACE".

In this case we are referring to the ACE satellite, not the ACE-MAESTRO instrument. We've clarified this by saying "...instrument on board the ACE satellite."

Pages 16289/90: For MAESTRO and SPS: for further validation campaigns, it probably would be well worth looking into some temperature stabilisation for either at least the PDA or preferably for the whole instrument. Although an initial expense, this should well pay off re data improvement. Page 16292, line 6: "The NO2 is near the limits of detection of SPS and MAESTRO". There might be other issues as well, but (as mentioned above) temperature stabilisation might well improve the situation substantially.

We agree that temperature stabilisation may improve the data. The instruments were installed inside a commercial freezer during the MANTRA 2004 campaign to attempt to stabilise the temperature of the detectors. This installation proved to be problematic for other reasons.

Page 16292, line 13: I assume there were no box model results for 80N and that is why
the NO2 profiles for 75N is used?

Correct. The model can be run at either 75N or 85N. The results at 75N are more reliable, and so we have chosen this latitude.

Page 16293 Table 2 Figure 4: Is the 5% (for ozone) and 20% (NO2) error assumption really realistic for MAESTRO and SPS, given the noisier spectra, etc.?

The noisiness of the spectra is considered in the "random noise" component of the total error, given in Table 2.

Page 16295, lines 7-9: Sounds somewhat confusing that "the data measured at SZA > 92 is unreliable because the thermal noise increases as the detector gets warmer". Not sure what you want to say here exactly. It is obvious that the thermal noise increases when the detector gets warmer but then the temperature should decrease with increasing SZA. Or do you mean to say that in general the quality of the spectra measured at high SZA gets worse while the campaign progresses and ambient temperatures are increasing?

SAOZ is operated inside PEARL in a hatch looking through a plexiglas window. The temperature in the hatch is fairly constant through the day. Through the campaign it ranges from 25 to 30 degrees C, which is a bit warmer than ideal. The consistently high temperature has more of an effect on the spectra taken at high SZA because the amount of sunlight is small, and the dark signal (including the thermal noise) makes up a greater proportion of the total signal. We've clarified this point in the text:

"This divergence is due to the consistently warm temperatures (25-30C) inside the viewing hatch of the instrument. The data beyond 92 degrees is unreliable due to the dark signal (which includes the thermal noise) making up a larger percentage of the total signal."

Page 16302, lines 6-7 Figure 13: Either indicate 8 March in Figure 13 or add day number to text.
We've added the day number (67) in the text.

Page 16307, lines 14-15: Why was the smaller altitude range (22-40 km instead of 13-58 km) chosen for the NO2 partial columns?

The Kar et al. paper recommends the 22-40 km region as being suitable for scientific analysis for ACE-MAESTRO. We take the smaller region for both satellite instruments to allow for direct comparisons between the partial columns. This is described when discussing the ozone comparisons, which now directly proceeds the NO2 discussion.

Table 1: Why did you use 400-450 nm for the NO2 analysis for MAESTRO and SPS while 400-550 nm was used for the other 2 instruments. Again because of the noisier spectra?

For MAESTRO and SPS we found it necessary to perform a separate calibration for the ozone and NO2 regions. For SAOZ and UT-GBS one calibration was sufficient for both species. This is likely due to the noisier spectra.

We’ve added this discussion to Section 4.1.

Figure 1: Text is rather small and hard to read.

We hope that the figure will be larger in the ACP format of the paper, and that the text will be more legible.

Throughout the manuscript, sometimes "Figure" is used, sometimes "Fig.", shouldn’t that be consistent?

The ACP(D) style guide requires "Fig." if the reference comes in the middle of a sentence and "Figure" if it begins the sentence.