Thank you very much for your thorough and thoughtful comments. We believe that in addressing these comments, the manuscript will improve significantly.

General comments:

In section 2, several simulation methods are given, which include a multitude of parameters.

In section 3, the results of these simulations are presented. Here, the references to simulation parameters are given by (only) partly repeating these from section 2. At initial reading, this causes difficulty and a lot of back-and-forth reading in relating the results to the simulation parameters. And I still have to guess.

I strongly suggest to make this more structured, by clearly itemizing in section 2, like "simulation A: ...", "simulation B: ..."

and then referring to these cases in section 3

We will delineate the different simulations more clearly by adding a table to section 2 that gives the key parameters (NO2 amount, solar zenith angle/time, scene type, polarization sensitivity, orientation, and locations) for each of the four main simulations. We will then clearly reference these in section 3 in the text and figure captions.

It is shown in the paper that instrument polarisation sensitivity mainly affects the AMF retrieval, not the NO2 slant columns.

In this study, the polarisation sensitivity enters NO2 vertical column through the retrieval of cloud fraction.

But in operational retrievals, also the cloud altitude and(or) cloud optical thickness must be derived.

Especially for cloud retrieval that uses information from polarised radiance (e.g. rotational raman scattering or deep absorption bands of O2) it may be expected that instrument polarisation plays a role. It is understood that the cloud retrieval algorithm for an instrument in initial development is TBD and simulations are premature. Nevertheless this should be mentioned explicitly.

We added the following: “Our simplified retrieval approach may have neglected factors used in operational retrievals that could be affected by instrument PS and

contribute to additional retrieval errors related to estimates of aerosols, surface reflectance, and cloud parameters."

Specific comments:

Figure 2: Surface spectra are shown here from ~400 to 3000+ nm. But the NO2 retrieval window is 420-455 nm and from these figures it is impossible to see any spectral structure there. Please reduce the spectral range of the figure and comment on the surface spectral resolution.

We have added a plot to Figure 2 showing the spectra in the 420-455 nm range.

lines 160-175: here the description becomes confusing/sloppy:

AMF in line 163 seems to be the height-dependent box-AMF which in Eq.(10) should be written as function of *z*.

I assume you mean AMF in line 165.

Yes, we will clarify and specify box AMF

What is the relation between AMF_tot in Eq.(7) and AMF_total in Eq.(10) ? Meant is probably that Eq.(7) in incorporated in Eq.(10). I suggest to write this out in Eq.(10).

In order to explain the alpha in Eq.(10) I suggest to say beforehand that the formulation follows Kuhlmann 2015.

In line 175, should not ∂AMF_tot be ∂AMF_total, and is its only dependence on PS through Eq.(8) ? Please rewrite to make that explicit.

We will clarify this section by being more explicit and consistent in our AMF terminology. We have changed the equation 9-11:

- Eq. 9: AMF_{total} will replace AMF_{tot}
- Eq 10: The box AMF is shown as a function of z (and referenced in the text), and $AMF_{clr/cld}$ will replace AMF_{total}.
- Eq. 11: $\partial(AMF_{total})$ will replace $\partial(AMF_{tot})$

section 2.2.2:

were the same aerosol parameters as for clear sky used? or no aerosol at all?

There are no aerosols for the cloudy scenes. We will add the bolded text to: " we model Lclr as a Lambertian surface (opaque) with surface reflectivity 0.80 at the effective cloud pressure, assumed here to be equivalent to a cloud at 2 km. **Aerosols are not considered for the cloudy scenes, since they would have a negligible impact; the clouds would lie above the tropospheric NO2 and aerosol layer."**

it is tacitly assumed that cloudy pixels have unpolarised radiance. Please mention/explain this explicitly.

This is not the case in general. The confusion may be arising from the simplified illustration (Fig. 3) where the polarization response is not being applied to the cloudy scene. This is only true when we start from a clear scene and look at the how the deviation in radiance form the PS changes how this affect the interpretation of the cloud fraction. When we start with an assumed non-zero cloud fraction, the output of the RTM includes a combination of the simulated Stokes from the clear and cloudy scenes (when for instance, we used the GEOS-5 profiles and cloud fractions).

We will add the bold text in the following to clarify this point: "This deviates from the illustration in Fig. 1 (top left), where instead of a clear scene, a mixture of cloudy and clear scene according to the GEOS-5 cloud fraction value is used**, thereby accounting for the radiance polarization state of both clear and cloudy scenes in generating the NO2 retrieval errors"**

line 208: "other retrieval techniques that do not use a spectral fitting approach" should be

"other retrieval techniques that do not use a polynomial correction term in the spectral fitting approach"

Will change.

Figure 5b. Why is the standard deviation over Water so much smaller than over Land? Usually the reflectances over water are smaller so S/N should be worse, unless you force S/N to be constant (as suggested in the text). Or is something else the case like an aerosol effect or a spectral surface effect? Please explain.

The water scene had lower retrieval errors for this location. This is not always the case, and the differences in SNR do not always drive the retrieval errors. We have not isolated the cause, but it is likely a combination of SNR, view geometry, and aerosol models that contribute to the retrieval errors.

We will change, " The errors are driven by a combination of the noise and are similar for all scene types" to "The errors are driven **by a combination of the SNR, view/solar geometry, surface reflectance spectrum, and aerosol model** and are similar for all scene types."

section 3.2 . Confusing: which simulation from section 2.2.2 was used to generate Figure 7 and which one for Figure 9 ?

Please refer to the previous comment. Adding the table and references to it will hopefully clarify.

See my general comment.

Figure 7 has fixed surface type thus seems to be the second simulation from section 2.2.2. Figure 9 uses GEOS-5 data thus also seems to be the second simulation ??

The simulation identifier will be added to the caption

What means " water, rural, urban scene covers CONUS [...] for each surface? use for each grid cell the most abundant type?
What means "with a fixed scene type over the CONUS grid" ? use 1 type for all grid cells?

Please provide a bit less condensed description.

Added a sentence to 2.2.2 in bold: "Simulation B quantifies the retrieval impact of scene type, water, rural, and urban scene, over CONUS for a constant reference NO2 profile. **The scene types are the same as defined in Table A and are assigned to all pixels in CONUS for each run**. "

line 220: parenthesis typo in Zoogman reference

Will fix.

Retrieval results for Figure 6: Is it correct that these simulations were done with a small (<0.04) cloud fraction? "Cloud radiance fraction" refers to the retrieved result

(doesn't it ?). Please specify which cloud fraction was used in the forward simulation for this figure.

Will add the bolded text to the description of simulation "The simulation ran using 70° solar zenith angle and water scene for all pixels and an instrument PS of 5%, m01 = −0.05, vertical orientation and m02 = 0.05, 45◦ orientation, **and an initial cloud fraction of zero**." Will also add "**retrieved** cloud fraction" to refer to the result after considering PS effects.

Figure 8: data for NO2 amounts 0f 5.0E+15 and 8.6E+15 are difficult to read in this figure. Is the relative error ("percent error") approximately equal for all three NO2 amounts? Please adjust figure or mention in the text.

The % error will be added in the text. It is non-linear with NO2 amount. We will add: "The dependence on NO2 amount is non-linear; for instance, at 5 % PS for the Seattle, evening case, the retrieval errors for increasing amounts are 0.22 %, 2.6 %, and 6.6 %, respectively.

Retrieval results for Figure 9: are these retrievals with "fixed scene type" as suggested for the scenario with GEOS profiles in line 181? That would not be very realistic. If water is used for the extreme East/West (Atlantic/Pacific) why are the errors so much smaller than in fig 7c? Your text says "The higher cloud fraction decrease the retrieval errors" but also clear scenes at high solar zenith angle have much smaller errors. Is this because of the NO2 amounts? It would be useful to show a figure with NO2 input column.

Yes. Water was used for all pixels as clarified (see above comment for section 2.2.2). Although this isn't too realistic, the lower reflectance types gives the largest values. We used this type, because we are more interested in upper bounds for requirements considerations. We will add the NO2 amount per location in Figure 9 and the bolded text in the following: "**As a result of the cloud fraction and lower NO2 amount**, the maximum NO2 errors found were 0.03 $\times 10^{15}$molecules/cm^2 for this day —a negligible value when compared to the TEMPO precision requirement."