Review of the paper titled “**A comparative analysis of IV nadir-backscatter and infrared limb-emission ozone data assimilation**” by **Dragani**.

In this paper, the author investigates the impact of the addition of a nadir and a limb ozone profile data set, produced by ESA's O3-CCI, on the ozone fields in the ECMWF IFS Data Assimilation System. The global distribution of ozone and the change in uncertainty are studied and compared to a reference without additional ozone information. The two new data sets are also compared to a control run, and to external reference data (e.g.: MLS and ozone sonde).
The author demonstrates that both the nadir and the limb ozone profile data sets can improve the assimilation, and makes a case for more limb measurements when most of the satellite instruments to be launched in the near future are of the nadir class.

General Comments

In this paper I see the name of the satellite instrument and the retrieved ozone product(s) being used interchangeably a lot when they should be kept separate in my opinion. For example: P4, L1-3: In these sentences the two are starting to get intermixed.
While the instrument is unique, there are several implementations of retrieval algorithms for a particular satellite instrument's data. For example, for GOME-2 the ozone retrieval algorithms / products known are Miles et al (O3-CCI), Cai et al, and Hassinen et al (O3MSAF). The retrieval algorithms may all have different behaviour, which makes it harder to make general statements like 'GOME-2 data is....'. The same holds for MIPAS, the instrument (and L1 data) is not the same as the ozone product coming out of a retrieval algorithm. Please check the manuscript, and identify where you really mean the instrument, and where you refer to the ozone retrieval product.

Specific Comments

P1, L23: You mention the warming/cooling of the air in the atmosphere, and then mention its (long term) effect on climate. A warming / cooling of air has a more immediate effect on the atmosphere: a temperature difference leads to a density difference, which leads to a pressure difference, which in turn leads to flow of air. In that way, the global ozone distribution can affect the dynamics of the atmosphere.

P4, L12: Do you mean resolution, or sampling? The sampling of the instrument is usually defined as the distance between detector pixels (in nm), but the resolution of the instrument is also affected by the width of the instrument's slit function (which may be wider, and span more than one detector pixels).

P7, L22-25: The author starts off with mentioning that O3-CCI/GOME-2 has the largest difference in the 4 month period. While this is true in the beginning, it would be more insightful for the reader if the discussion on the differences would be split into the first two months and the last two months, as is the case in the later sentences where the O3-CCI/MIPAS differences are split into Jul/Aug

and Sept/Oct. Given that the behaviour of the difference changes with time I feel that giving a range of the average difference is more representative than a single value over the four month period.

P8, L8-11: Using the larger provided uncertainty to compensate for the fact that vertical correlations (by means of Averaging Kernels (AK's) and covariance matrices from the ozone retrieval) are not used in the assimilation systems is risky. Can the author give an estimate whether the larger provided retrieval uncertainty is similar in the value and the sign of the vertical correlations?

P8, L18: The author states that no corrections are applied for GOME-2 O3 nadir profile data above 5hPa but does not mention what kind of corrections are applied below. It leaves the reader in doubt on what happens. Only at P8-L32-34, at the end of the paragraph that discusses MIPAS, the reader finally finds that no corrections are applied to either the nadir or limb retrieval product. Please make this clear in the paragraph that discusses O3-CCI/GOME-2.

P9, L5-6: The author mentions that the behaviour in the tropical region is different for O3-CCI/GOME-2 than for O3-CCI/MIPAS. From figure 4 it is clear that the largest differences seem to coincide with the ITCZ, which is a clouded region in the tropics with high cloud tops. In P5-L9+10 the author states that the MIPAS data has been carefully screened for clouds, while I see no such statement for the O3-CCI/GOME-2 data. This could explain the large differences, when the nadir ozone profile retrievals are affected by ozone ghost columns, as the clouds block the observation of ozone below the cloud top. Would it be possible to investigate the effect of removal of pixels with a large cloud fraction from the O3-CCI/G2 data set on the global distribution of differences?

P10, L19: As far as I understand the comparison of two satellite retrievals, both instruments should be brought to a common grid and the AK's should be cross-applied. See Calisesi, et al (2005), Regridding of remote soundings: Formulation and application to ozone profile comparison, J. Geophys. Res., doi:10.1029/2005JD006122. Would this be a feasible approach for this study? For comparisons using reference data with a very high vertical resolution, such as ozone sondes, the transformation in Calisesi is not required because the ozone sonde's 'averaging kernel' peaks only near the measurement altitude (as it is a very localised measurement of the air it passes through).

P11, L29-31: Both instruments show reductions of the standard deviation (Fig 9). The one from O3-CCI/GOME-2 occurs over a wider vertical range than the one from O3-CCI/MIPAS, wheras the latter seems to have stronger localised reductions. Which of the two would be preferable for the assimilation as a whole and why?

General question

The author demonstrates that the comparison with MLS and ozone sondes improves when GOME-2 and MIPAS based ozone profiles are assimilated, but if it is not too far out of scope of this paper, it would be interesting to get an indication of the change in the skill of the IFS in general as a result of the

assimilation of the additional ozone input. E.g.: the effect on wind vectors or temperature.

Typographical comments

P2, L2: signature → … the **signing** of an international treaty … [signature is a noun, signing is the activity].

P3, L5: greatly → very

P4, L28: **The** CCI …

P5, L10: verified? You may mean 'present'.

P5, L32: satellite**s** (plural).

P6, L16: “An example of **a** background error profile and **a** vertical correlation…”

P6, L18: Introduction of acronym TCO3 without prior explanation (also not in table 1).

P12, L24: **is** equivalent

P12, L25: what → why

Figure 12: The plots are small and the black and blue are sometimes hard to distinguish with this line width. Would it be possible to provide larger plots with thicker line? One could try a 2-1-2 panel ordering instead of the current 3-2.

References:
Miles et al (2015): double doi
Munro et al (2006): Please check initials of Munro, it seems that there are spurious letters, as the other reference has an 'R.' only.