Interactive comment on "Post-processing to remove residual clouds from aerosol optical depth retrieved using the Advanced Along Track Scanning Radiometer" *by* Larisa Sogacheva et al.

Anonymous Referee #2

General comments

This paper presents an update to the post-retrieval cloud contamination method for satellite aerosol retrievals (and the ADV/ASV AATSR algorithm in particular) presented previously by Kolmonen et al. (2015). This update essentially involves a simple threshold test on the proportion of high AOD values retrieved in a region, with high-AOD regions not having the Kolmonen tests applied.

This rather simple test, although it has a significant impact on the resulting ADV/ASV aerosol product, is not particularly interesting in a scientific or retrieval theory sense. The paper is essentially an addendum to the previous one and one is left wondering

how a 13 page paper has been spun-up from it. Indeed, the reader has to plow through about 8 pages of reiteration of already published work before reaching the new work presented by this paper. I would like to see a significantly streamlined and shorter paper, focusing on the changes made from the methods and products described by Kolmonen et al. and showing some summary plots of the resulting improvement to the ADV/ASV product.

Thus my recommendation is that major revisions are required before this work can be published. This clearly puts me at odds with the already submitted review, so I have provided detailed suggestions and corrections for the paper in its current form below.

Specific corrections
In the introduction the authors describe what might be described as traditional cloud masking - a series of empirical tests based on thresholds in radiance, radiance ratios or spatial variability - as though this is the only approach used for cloud detection. This isn't true, as there exist other approaches, such as neural-network (or other machine learning techniques) or statistical approaches such as naive Bayesian fitting have also been used. The authors should acknowledge this fact, and, crucially note that these approaches also suffer from the same limitations as traditional threshold based techniques.

Additionally, the authors should mention that a lot of the difficulties encounter with cloud-masking are due to its inherently subjective nature - what is a perfectly adequate mask for one task may well provide far from ideal to for another (e.g. a cloud mask used for cloud retrieval needs to be sure it IS looking at cloud, while one for aerosol retrieval needs to be sure it ISN'T looking at cloud).

I have a problem with the use of the word "plume" in describing regions of high AOD from section 5 onwards. An aerosol plume is a clearly defined aerosol stream with a identifiable source, such as a smoke plume from a fire. The authors use the word to mean a region with elevated AOD, which is potentially confusing. Please use a more appropriate name, like high-AOD region for instance.

Abstract
- P01L08: Replace "much stronger than" with "much more strongly than".
- P01L10: Not all cloud masking techniques rely on a test of empirical tests - for instance neural-net or naive Bayesian approaches. Thus, replace "consisting of several tests. However" with "but".
- P01L13: Replace "shows such areas with locally enhanced AOD values" with "produces locally enhanced AOD values in such areas".
Introduction
- P02L12: The Kaufmann et al. and Zhang et al. papers are both relevant to a now quite outdated MODIS product version (although it isn't clear from the papers themselves which "collection" they used). The authors might consider providing more up-to-date references, or at least acknowledging that MODIS products - including cloud flagging - have improved markedly over the past decade.
- P02L22-L25: This section seems to be providing a list of attributes for an ideal aerosol remote sensing instrument rather than one for cloud-masking... I think the statements about how current instruments sadly fall short of the idea for aerosol retrieval are superfluous. Just note that the cloud detection approaches for different instruments make use of the differing capabilities of those instruments.
- P02L32: Replace "refer to other methods" with "use other methods"
- P02L32: Sentence starting "Sensors with narrow spectral channels..." is irrelevant. Remove.
Section 2
- P04L16: Replace "The AATSR is a dual view instrument" with "The ATSRs are dual

view instruments".

- P04L20: Pedantically, the 1.6 μm band is near-IR, not mid-IR.

Section 3

- P05L02: Why is the 0.865 μm band only used over the ocean?

Section 4

- P05L16: The term "native pixels" is inaccurate, as the standard AATSR level 1 files (which I presume ADV uses as its input data) are regridded from the instruments native sampling to a regular 1x1 km grid using nearest-neighbour sampling. This is a particularly important distinction when talking about the forward view, where multiple 1 km pixels may actually be filled by a single instrument pixel.

- P05L24: I am confused by the reference to Press et al. (1992) here - how is "Numerical Recipes" a reference for the automated determination of cloud masking thresholds? The same question applies to P06L17.

- P05L25-P06L22 (Section 4.1): I think a bulleted or numbered list of the four tests would be clearer.

- P05L26: "Most clouds are cooler than the underlying surface...". Also, can the authors provide an indication of how sensitive this gross cloud test is? How big does the temperature difference have to be for the pixel to be flagged? What happens, for instance, over mountainous regions, where the surface can be expected to produce large variations in brightness temperature?

- P05L29: Replace "below threshold" with "below this threshold".

- P06L01: Is the test based on the difference in absorption coefficient of water vapour from 11 - 12 μm, or is it the emissivity of ice at these two wavelengths? I think you'll find it is the latter.

- P06L03: "$\Delta T = T11 - T12$ is usually positive..."

- P06L20: "reflectance at 0.865 μm is higher than at 0.659 μm over vegetated areas".

- P06L23: Replace "native" with "Level 1B" (see comment for P05L16).

- P07L08: Replace "For the dual view instrument" with "For the imagery provided by a dual-view instrument"

- P07L09: Replace the sentence beginning "The reason for the stereo effect..." with "In (A)ATSR Level 1B data the forward and nadir data are nominally collocated at a standard surface ellipsoid corresponding to sea level". Also replace "In case of a cloud, the forward..." with "In the case of a cloud located above a given surface pixel, the forward..."

- P07L10: Replace "while the nadir view already sees the cloud approximately 2 minutes later" with "while the nadir view sees a fully cloudy pixel".

- P07L11: Remove the two sentences starting "In the AATSR data..." and "Hence the foward view..." The forward and nadir views of AATSR are NOT collocated in time, they are spatially collocated at sea-level!

- P07L17: "depending on the cloud shape and height". Also, "both nadir and forward views are used in ADV over land".

- P07L21: Reword sentence to reference Fig.3: "The AOD retrieved from the same test scene used in Fig. 1 and Fig. 2 is shown in Fig. 3".

Section 5

- P08, first paragraph: The term pixel is now referring to the 0.1 degree retrieval pixel, not the 1 km Level 1 pixel, right? Please explicitly state this, or use a different word/phrase for the level 2 retrieval pixels.

- P09L05-L10: The description of how a region of high-AOD (a "plume") is defined is rather confusing. There appears to be a single test applied - what fraction of retrieval pixels produce an AOD > 0.6, but the authors talk about multiple thresholds and a multi-step process. I think it would be clearer to state up-front that high-AOD regions are defined as areas where >40- P09L13: Insert text "The thresholds for the number of cloud-free pixels..."

Section 6

- P11L05: I find the statement that level 1.5 AERONET data hasn't been compared against because "validation must be done with the best quality data" a bit of a cop-out. If the authors feel that using level 1.5 data might provide better coverage of high-AOD events, then they should try it, even if it cannot be considered a primary validation.

- P11L06: Replace "retrieved AOD and AOD results..." with "retrieved AOD and AERONET results..."
- P11L18-L20: The sentences beginning "There is a discussion..." and "The linear approximation of..." are superfluous. Remove them.
- P11L22: Replace "Instead we present the mean AOD and AOD standard deviation in different AOD bins (fig. 8)." with "The binned AOD mean and standard deviation are also indicated in Fig. 8."

Section 7
- P11L28: Replace "entails changes in..." with "results in changes in..."
- P12L18: Delete the sentence beginning "Regions of interest shown in Fig. 12..." (the caption of the figure provides this information).

Section 8
- P13L06-L09: This paragraph ("Yearly AOD time series over land...") and the associated Fig. 14 are redundant. I suggest they both be removed.
- P13L10-L14: Conversely, why is there no figure showing the comparison with MODIS Collection 5? Furthermore, why are the authors not using more up-to-date MODIS data (Collection 5.1 or 6)? Please add a figure, or remove this paragraph.

Tables and Figures
- Table 2: In the caption, note that nnnnn corresponds to ENVISAT orbit numbers.
- Table 3: The dotted style of the cell border lines in this table make it hard to read.
- Figure 1: In the caption, define which bands have been used to produce the RGB image. Also, note that it is vegetation which appears read in this imagery, not land per se. Finally, the correct grammar is "ocean is coloured black", not "ocean is coloured in black".
- Figure 8: The individual plots in this figure are too small to clearly make out the binned-averaged AOD values and their associated error bars. Perhaps split the figure into separate figures for each region? Also, extend the 1:1 line to span the full plotting range (0-4, rather than 0-3).
- Figure 10: The top two plots are multi-annual means, not yearly means. Also, define

the seasons in the caption; is winter DJF?

- Figure 11: Is this figure necessary? It doesn't convey anything not obvious from Fig 10.

- Figure 13: This figure is far too busy and confusing. It is extremely difficult to distinguish which coloured line corresponds to which region. Also, the caption needs to define exactly what is plotted in each panel.

- Figure 14: I think this figure is superfluous (see comment for P13L06).