Replies to Reviewer #2's comments

First of all the authors gratefully acknowledge the reviewer for his/her comments and suggestions. The reply to each reviewer's comment is given in blue below the comment.

Interactive comment on "The Advanced Infra-Red Water Vapour Estimator (AIRWAVE) version 2: algorithm evolution, dataset description and performance improvements" by Castelli et al., 2018
Anonymous Referee #2

General comments

This paper describes a new version of the AIRWAVE algorithm and the retrieved Total Column Water Vapour (TCWV) product from the Along-Track Scanning Radiometer (ATSR) instruments. TCWV is obtained from the thermal infrared channels at 10.8 and 12 μm, the nadir and oblique views, and over sea surfaces. Validations are carried out based on aggregated products at 0.25 deg resolution and worldwide.
Overall, the reading of this paper leaves me in two minds. On the one hand, the value is principle very high due to the legacy of the very long time series of dataset provided by the ATSR series and the importance of deriving TCWV. The retrieval approach seems interesting too with the decomposition of the inverse model into a linear analytical equation. And the apparent evaluation results given at the end seem encouraging for this product and associated retrieval approach. But on the other hand, the added value and independency of this new manuscript w.r.t the requirements of the AMT journal and work already done and presented in the pre-existing studies, on which this manuscript relies, seem relatively little. I cannot at this stage recommend publication in AMT, and would like first to encourage the authors to address the major comments listed below:

- As emphasized by Referee #1, figures and layout of this manuscript are almost identical to the paper of Papandrea et al. (2018). The main difference is the evaluation of the v1 dataset, while here the focus is on v2. Apart of the generation of v2 (and associated new simulations), the present paper does not provide new works and innovations compared to the mentioned previous works.

As replied to reviewer #1, it is true that the figures have the same outline of Papandrea et al 2018. This was done to ease the comparison between the two dataset versions. However, the content of the figures is totally different! and also the layout of the paper is not the same. In this paper we present the new version of the algorithm with all the new calculations used to improve the retrieval performances. Then we show the performances of the new dataset against the AIRWAVEv1 and we present the climatology from AIRWAVEv2 dataset. In addition, we provide a quantification of major error sources. Due to the entity of the improvements of the TCWV AIRWAVEv2 dataset we think that the paper presents new work with respect to the previous ones.

- In several parts of this manuscript, the writing is a bit too qualitative, and even sometimes a bit ambiguous. Efforts shall be made to add more quantitative elements.

Done. Following the reviewer's suggestions, we add more quantification of the results and clarified the points raised by the reviewer.

- The reader is sort out invited to check the improvements between AIRWAVE v1 and v2 by himself with the few lines given on Page 9. This is a bit unusual. A direct and explicit inter- comparison shall be provided here, with an explicit estimation of the bias reduction and precision improvements worldwide, per areas, and for different conditions.

Done. We report the results of the validation of AIRWAVEv1 in Table 2 together with AIRWAVEv2. Then we add a discussion on that and a direct comparison in Section 4.

- One of the expected improvements is the across-track variability. However, this element is not quantified in terms of uncertainty reduction on the TCWV. Same regarding the latitude dependency. Also please clarify sometimes if you are talking about bias, precision, overall uncertainty, etc...

In the revised version of the paper we clarify when we talk about bias or precision. Regarding the across track variability a more extensive reply is given below. In Appendix A we give an estimate of the accuracy of the new parameters using simulated values. When applied to real data, due to the nature of the algorithm and to the used interpolations, is not easy to decouple the single effects on the retrieval. However, in the paper we report the results (in terms of bias and standard deviation) of the validation against other instruments also for latitude regions in order to highlight the impact of the new algorithm in different latitude bands.

- To warrant a new publication and invite the use of this new dataset, I think this manuscript should go beyond the previous ones by providing more validation exercises with new additional TCWV dataset instead of repeating similar work: from other satellite instruments: e.g. MODIS (Diedrich et al., 2015), MERIS (e.g. Lindstrot et al., 2012), ENVISAT Radiometer, MSG SEVIRI, and / or ground-based sensors (AERONET, GNSS, or other radiosondes), and / or ECMWF reanalysis. The approach proposed by AIRWAVE shall critically be analysed and compared with more "classically achieved" with other thermal infrared sensors to be in line with the ambitious title. Furthermore, what about the different sensitivities to the atmospheric layers of the different spectral ranges (i.e. thermal infrared and microwave). Does it allow, prevent, or limit in a certain range the comparison of the associated total columns?

The main scopes of this paper is to describe the algorithm evolution, the dataset and performance improvements. For these reasons we repeat the validation exercise using the same validation dataset used for AIRWAVEv1. Furthermore, the validation is performed using satellite microwave data and radiosondes data, thus two very different datasets (platforms, sensitivity …) allowing to evaluate the AIRWAVEv2 performances under two very different point of view. In future, additional validation exercises will be performed (e.g. with MWR on ENVISAT), however this is out of the scope of the present paper.
Regarding the different sensitivities of TIR and microwave, in the validation exercise we compare SSM/I data in coincidence with AIRWAVEv2 ones. Since AIRWAVE is applicable only to clear sky measurements this is a method to filter out SSM/I "cloudy" TCWV and thus to avoid biases due to different sensitivity related to the used spectral range. When comparing to radiosondes, the small bias we found demonstrate that AIRWAVE TCWV are sensitive also to low atmospheric levels.

In the revised version of the paper we added a sentence about the different sensitivity of TIR, microwave and radiosondes reported above.

- The evolutions in the AIRWAVE equations in Sect. 2.1., i.e. "latitude dependencies and across-track variations", do not explicitly appear. It is quite hard for the reader to be able to make the link between these, the given variables, and their physical meanings. The authors shall better help the reader to establish these connections. Furthermore, I am a bit surprise that across- track and along-track viewing angles were not already considered in the previous version. And adding this in the update seems quite natural I think. As far as I can see, every atmospheric retrieval approach based on satellite sensors always needs to consider the sensor geometries to characterize accurately the average light path followed by the detected photons. Given how the air mass thickness varies with the geometry, the consideration of the viewing angles does not sound as very innovative to me, but rather quite natural. Other (precipitable) water retrieval from ATSR2, e.g. Li et al., 2003, although over land, Ren et al. (2015) or even from SEVIRI (e.g. Sobrino et al., 2007) already considered the variations of zenith angle at the surface for both nadir and forward views.

In the revised version of the paper we follow the reviewer's suggestion reported also in the specific comments and tried to clarify the quantities in the equations, their link with viewing geometries and physical quantities.

Regarding the across track angles: possibly there is a misunderstanding here. In Casado et al., 2016 : "To account for the variability of the nadir view angle (0–25°), the TCWV evaluated through Eq. (29) are corrected using an empirical correction factor, dependent on the across track index position of the ATSR measurements. This across-track correction has been evaluated through radiative transfer simulation and provides a sufficiently accurate but extremely fast solution (in terms of CPU efficiency) to this problem. In the next version of the AIRWAVE algorithm, the retrieval parameters will be extracted from pre-calculated look-up tables, in which the angular dependence will be considered." And what we present in this paper is exactly the calculation of angular dependent retrieval parameters.
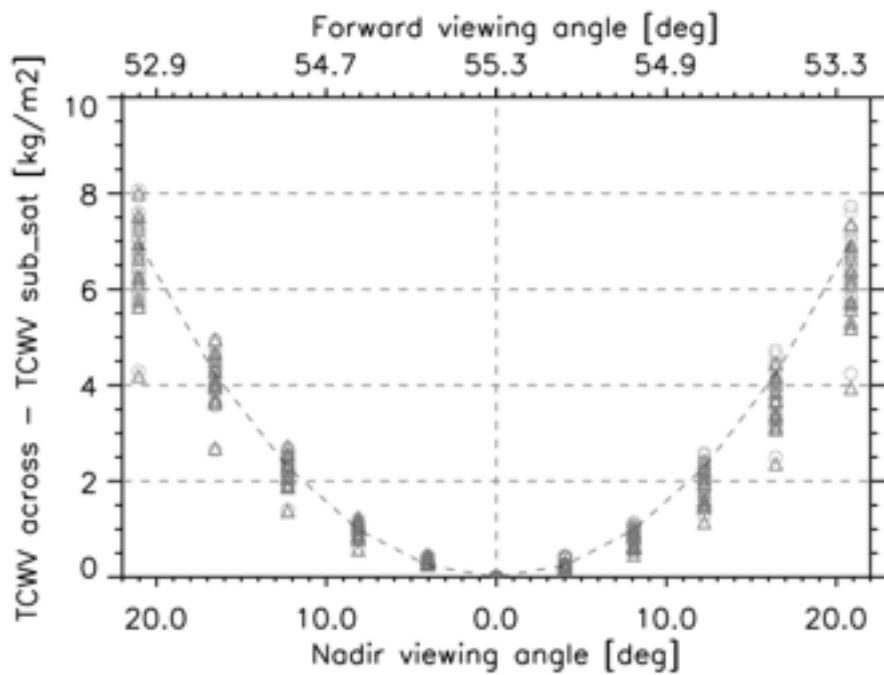

Figure 0: Differences between TCWV calculated from across pixels and TCWV calculated sub-satellite track as a function of across track position calculated for different atmospheric scenarios and for AATSR. Black dotted line corresponds to AIRWAVEv1 parametrization.

Figure 0 shows the absolute difference between TCWV from across track pixels and sub-satellite pixels calculated for AATSR in different atmospheric scenarios (tropical and mid-latitude) from synthetic measurements. To produce these retrievals, we simulated ATSR(s) radiances at the eleven tie points for the nadir and forward views considering also the surface emissivity variations with viewing angles. The atmospheric scenario, the TCWV and the SST were kept constant and they were exactly the same used for the sub-satellite track case. Then we retrieved the TCWV for each of the 11 couples of BTs and computed the difference with respect to the value obtained at the sub-satellite track position (that coincide with the TCWV reference value). The dotted line in Fig. 0 mimics the correction term adopted within the AIRWAVEv1 algorithm. As can be seen the dotted line well mimic the general behavior of the across track TCWV dependence. This is confirmed by the comparison made between AIRWAVEv1 TCWV and SSMI or ECMWF TCWV (see Casadio et al., 2016): on average we did not found any across-track bias, thus confirming the general validity of this correction. However, as can be seen in Fig. 0, in specific cases this approximation can generate differences depending on the used atmospheric scenario. Furthermore, a slight asymmetry with respect to the along track position is also visible. This behavior is expected as the ATSR instruments are tilted of about 4 respect to the flying direction of the satellite. In order to properly account for these variations, in AIRWAVEv2 we replace the a-posteriori correction by directly calculating the retrieval parameters for each of the 11 tie points of the nadir and forward swaths and interpolating the results at each position of the ground pixels. We also use in the solving equations the

correct emissivity values, in order to properly take into account the angular dependence of the sea surface emissivity.

In the revised version of the paper we add some details to clarify this (see replies to specific comments # 8,18 and 23).

Specific comments
1. Abstract line 4: "performs the TCWV retrieval" => Please rephrase more direct e.g. "retrieves TCWV"

Done.

2. Abstract line 5: "combining nadir and forward observation geometries": The verb "combine, here and elsewhere in the manuscript, sounds overstated for me. Both geometries are not considered strictly combined in a way it helps to retrieve 1 single TCWV value, But both views are considered individually and independently for deriving TCWV per observation geometry…

We remove "combining" in the abstract. Then we use "using" in other part (line 31 of the discussion version of the paper).

3. Abstract line 7: "almost no bias" => Please be more quantitative and less qualitative. What does it mean "no bias"? I do not know any retrieval with an exactly null bias.

We rephrase "very good agreement with almost no bias all over the ATSR missions, with the exception of the polar and the coastal region where AIRWAVE underestimate the TCWV amount." —> "very good agreement with an overall bias of 3% all over the ATSR missions. A large contribution to this bias comes from the polar and the coastal region where AIRWAVE underestimate the TCWV amount."

4. Abstract line 9: "these problems" => Which ones?

The underestimation of TCWV in coastal and polar areas. We change "to overcome these problems" with "to reduce the bias in these regions".

5. Abstract line 14: "significant improvements..." => Again, be quantitative. How much is the improvement overall? RMSE not defined previously. Clarify overall that AIRWAVE is over sea, no land!

Done. Now it is: "Results show significant improvements in both biases (from 0.72 to 0.02 kg/m2) and standard deviations (from 5.75 to 4.69 kg/m2 versus SSM/I)". We add "over sea" in the fourth abstract sentence: "The algorithm was used to produce a TCWV database over sea from the whole ATSR mission".

6. Introduction line 36: Why AIRWAVE v1 is only available aggregated at the coarse resolution of 2x2 deg, and not for all the individual retrievals?

AIRWAVE is included into the G-VAP at this resolution. Contacting the authors, the complete dataset can be obtained. Actually on the G-VAP only monthly means from 2003-2008 are available.

7. Introduction line 38 "good results" & line 42 "good quality": again, please be more specific and quantify. What the range of bias and precision for which you consider this is a good result?

Done. We added: "(average correlative bias of 0.72 kg/m2 vs SSM/I and 0.80 kg/m2 vs ARSA, below the 1 kg/m2 indicated in the GlobVapour project (Lindstrot et al., 2010))" after This exercise demonstrated a general good quality of AIRWAVEv1 dataset
In the PVR of the GlobVapour project report they use as goal 1kg/m2 for bias. Results from both AIWAVEv1 and AIRWAVEv2 are below this threshold.

8. Introduction line 44: "by accounting for latitudinal and angular variations of the retrieval": Not clear, retrieval estimations always vary with respect to geometry and latitude variation of the airmass. You probably mean the dependency of your forward & fitted model?

First of all we want to clarify that AIRWAVE does not make direct use of a forward model, but uses pre-computed parameters (retrieval parameters) to solve equation 13 of the revised version of the paper. In AIRWAVEv1 we used fixed parameters along the globe, in addition, "to account for the variability of the nadir view angle (0–25°), the TCWV evaluated through Eq. (29) are corrected using an empirical correction factor, dependent on the across track index position of the ATSR measurements."(Casadio et al., 2016). In AIRWAVEv2 we use latitudinal dependent retrieval parameters calculated for each of the 22 (11 nadir and 11 forward) tie angles. To clarify this,in the revised version of the paper we add :
"AIRWAVEv1 use fixed retrieval parameters along the globe and TCWV are corrected for viewing angles variability in nadir and slant by using an empirical correction factor." in line 39 after " For this reason, in the V1 algorithm several approximations were made."

9. Introduction line 48: "spread" => Do you mean precision? Or uncertainty? Or spread of the differences with respect to another dataset?

Spread of the differences with respect to another dataset. We modified the text accordingly.

10. Section 2 line 58 "average retrieval parameters": Which ones are you talking about here? Please be more specific. At this stage, the read haven't read it the equations in Sect. 2.1…

We reformulate:" AIRWAVEv1 makes use of retrieval parameters calculated though RTM simulations of tropical and mid-latitude atmospheric scenarios then averaged and used for the whole globe."

11. Section 2 line 66: again please clarify. Retrievals always vary with seasons and latitudes (water vapour properties). Do you mean that you explicitly consider the spatial and temporal variability of the some of the input parameters (e.g. H2O & temperature profiles)?

We mean that we did not use average retrieval parameters as done in AIRWAVEv1, instead we compute the retrieval parameters for 6 latitude bands and 4 months. Then during the retrieval these are used as look-up-tables. In the revised version of the paper we rephrased:
"Secondly, we computed the retrieval parameters taking into account both their seasonal and latitudinal dependence" —> "Secondly, we compute the retrieval parameters for different latitude bands and for four months that, in the retrieval, are used as look-up-tables"

12. Section 2 line 70 "We recall': Was never said earlier in this paper.

We removed " We recall here that".

13. Section 2.1 line 83 "F includes the atmospheric and surface radiance contribution". What is exactly F? This is not clear. Do you mean this is related to the temperature profile? Sea surface emissivity is already included later in the equation, so what is left? The temperature surface?

In the revised version of the paper we clarified this by explicitly indicating what is F. See also reply to reviewer #1.

14. Section 2.1 line 93 "relative effective absorption cross-section": what do relative and effective mean here? Do you mean the absorption cross-section integrated along the average light path (by opposition to the vertical atmospheric layers)?

With effective here we mean that the cross section is multiplied for the lambda value. This is the same definition used in Casadio et al., 2016. In the revised version of the paper we replaced "effective absorption cross section (sigma)" with "effective absorption cross section (lambda*sigma)"

15. Section 2.1 line 104 "We verified...": Where is it shown? Please support your claimed verifications with adequate figures to convince the reader. Quantify this linearity (e.g. high correlation coefficient value).

To clarify this we add a figure in the revised version of the paper and we modify the text as follows:
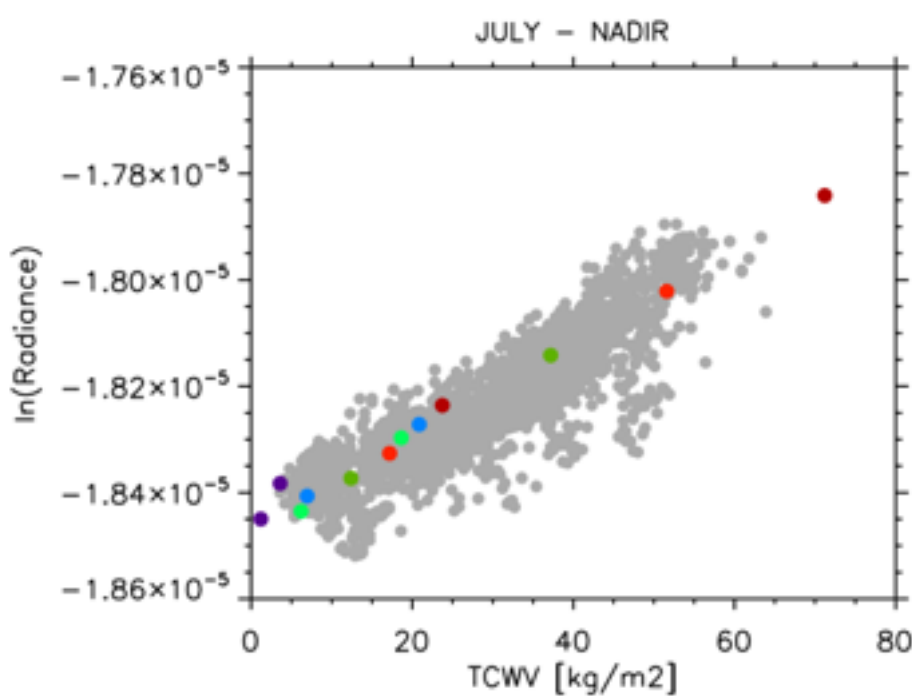

Figure 1: logarithm of radiance ratio at nadir as a function of TCWV in simulated atmospheric scenario. Grey dots represent the same quantity using real AATSR radiances and coincident SSMI TCWV for the along track measurements on the 5 and 6 August 2008.

In Fig. 1 the colored dots represent the values of the logarithm of radiance ratio in equation (1) as a function of the TCWV for the different atmospheric scenarios. We report only the values obtained for the sub satellite scans using the IG2 water profiles for the Summer season multiplied for 0.5 and 1.5. The different colors represent different latitude bands (going from red for tropical to blue for polar). The grey dots represent the radiance ratio calculated from along track AATSR measurements on the 5 and 6 of August 2008 aggregated at SSMI resolution (0.25° 0.25°). The value of TCWV associated to each AATSR sub-satellite measurement was obtained from coincident SSMI measurements. In order to minimise the impact of random error, only measurements with SSMI pixel coverage (calculated as the ratio between the actual and the maximum number of ATSR measurements that can be present into a SSMI pixel) greater than 10% were used for this exercise. Figure 1 shows that: a) the simulated radiances correctly reproduce the real measurement behavior; b) the relation between the radiances and the TCWV can be considered as linear.

Actually, we find in this case a correlation of 0.904 for real data and 0.92 for the simulated ones (p-value of 7.2630294e-05).

16. Section 2.1 line 110 "average of all the G values obtained with different water vapour content": Please clarify the series of values for H2O content that you used or the list of atmospheric conditions. Where do appear the zenith angles? They should already be, in theory, in G. J, emissivities, and F no?

We added: "(to vary the water vapour content we multiplied the water vapor profile for 0.5, 0.75, 1., 1.25, 1.5)"
after
" For this reason, for each atmospheric scenario the average of all the G values obtained with different water vapour content is used".
For the zenith angles the reviewer is right, we added additional information on this, see reply to comments 8, 18 and 23.

17. Section 2.1 line 117 "sort of effective water vapour cross section": Please reformulate more properly. Such a description is rather vague and ambiguous. What do you mean by effective here?

We remove the sentence.

18. Section 2.1 Eq.s 10 et al: Please be more specific by adding clearly all the angles in each equation: each view has a specific viewing angle. Furthermore, do you consider exactly the right zenith angle per pixel sensor? Or do you have a kind of average zenith angle for nadir and forward views? The Nadir view (cf. "NAD") cannot have only 1 zenith angle, or I miss something...

From equation 12 onward we explicit the angular dependence. We calculate all the parameters at the 11 tie point positions. It means that we calculate the equations for 11 nadir angles (from 0 to 21 degrees) and 11 Forward angles (from 53 to 55.3 °). Then we interpolate using each single pixel across track position.
To clarify this, in the paper we added: "Equations (11), (12), (13) were solved for the 11 couples of viewing angles corresponding to the tie points. The angles cover a range from 0 to 21 degrees in the NAD case and from 53 to 55 degrees in the FWD case." after "Equations (10), (12) and (13) are the solving equations used in AIRWAVEv2, while AIRWAVEv1 makes use of equations (11), (12) and (13). "

19. Section 2.1 line 136 "direct effect on the retrieval precision": And accuracy as well?

A preliminary assessment of the accuracy of the new method can be found in Appendix A, Section 5.1 using simulated radiances.

20. Section 2.2 line 173 "fixed wind speed": Why is it fixed? Why this value?

Due to wind speed variability, in both algorithm version, we prefer to fix the value of the wind speed and then treat this as an error source (see section 5.3). However, as can be noticed, this impact is negligible in equatorial regions, very low (0.7%) in midlatitude regions and is about 2.7 % in polar regions when the wind speed reaches the high value of 25 m/s. We choose this value (3 m/s) as a reference because is a reasonable trade-off between 1 m/s and 25 m/s that are extreme values available in the University of Edinburgh emissivity database. The 3 m/s is quite close to the average wind speed over the sea that is about 6 m/s (e.g. Monhan: "The Probability Distribution of Sea Surface Wind Speeds. Part I: Theory and SeaWinds Observations", Journal of Climate, 19, 2006. or data from Copernicus Marine environment monitoring centre data on 100km monthly L$ 2007-2012 Climatology Global Wind).

In the revised paper in section 2.2 we add: "Due to wind speed variability, in both algorithm version, we prefer to fix the value of the wind speed and then treat wind variations as an error source (see Sect 5.3)", before the end of the section.

21. Section 2.2 line 235 and further: "a good correlation is obtained against both datasets, as highlighted by the correlations and bias values reported in the same figure". Please don't let the reader thinking and looking by himself for these numbers. Report them adequately in the relevant paragraphs.

Following the reviewer's suggestion in this section we changed:
"Globally, a good correlation is obtained against both datasets, as highlighted by the correlations and bias values reported in the same figure."
with
"Globally, a good correlation is obtained against both datasets, as highlighted by the correlations (0.948 with SSM/I and 0.918 with ARSA) and bias values (0.02 +/- 4.79 kg/m2 with respect to SSM/I and 0.19 +/- 6.12 kg/m2 with respect to ARSA)."

and
"The improvement in the performances of the new dataset is clearly visible at all latitudes."
with
"The improvement in the performances of the new dataset is clearly visible at all latitudes, and in particular for regions with latitude higher than 45-50° where the negative values obtained with AIRWAVEv1 disappear. In addition, a significant reduction of the spread is highlighted."

Then we add in Table 2 the results of the validation of AIRWAVEv1 as extracted from Papandrea et al., 2018.

22. Section 2.2 line 260: How the new algorithm reduces the impact of the sensor noise? This is not explained before. Any evidence?

It is at the end of Section 2.1 we evaluate the impact of the sensor noise using the new equations.

23. Section 2.3 lines 184-185: I don't fully understand here. What do you interpolate exactly? I guess that now for each individual sensor pixel, you use the adequate zenith angle right?

As stated before, we calculate the retrieval parameters for the 11 tie points. Then using the across track position of each sensor pixel we interpolate the retrieval parameters calculated for the 11 tie points angles at the exact pixel position (and zenith angles).
To clarify we replaced:" In V2 we replaced the a-posteriori correction by directly calculating the retrieval parameters for each of the above described tie points of the nadir and forward swaths and interpolating the results at each position of the ground pixels."
with:
"In AIRWAVEv2 we replaced the a-posteriori correction: We calculated the retrieval parameters for each of the above described tie points of the nadir and forward swaths, then we obtained the parameters at the exact ground pixel position interpolating these values and using the ground pixel across track position."

24. Section 4 and appendix: Section 4 deserves more quantitative results with rigorous evaluation and validation of the AIRWAVE v2 to be in line with the title. More dataset, and more quantitative analyses per scan angle (since this is one of the claimed improvement). Also, how well does this approach achieve w.r.t to more classical approaches considered on thermal infrared sensors (e.g. Landsat, METEOSAT, etc...)? Moreover, I don't understand why Section 5 is here reported as an appendix. Some claims are written in the conclusion section, but were never reported before. I had to find out that some additional works were done in this appendix which is not obvious. Also what about the impact of the sea temperature surface? And H2o profile? Does it play a rule somewhere?

In Section 4 we added the results of direct comparison of AIRWAVEv1 and AIRWAVEv2 adding in table 2 the results of AIRWAVEv1 validation.

About the analysis per scan angles: as shown in Fig.0 above, the variations with scan angles depend also on season. AIRWAVEv2 approach is to select the parameters from look up tables through a multivariate interpolation on a 3-dimensional grid (trilinear). Decouple the effects of latitude and scan angle variation is thus not straightforward.

Regarding the validation with additional datasets, this is behind the scope of this paper. Actually, we present here the validation with both satellite data and radiosondes. The intrinsic nature of these two datasets and the AIRWAVE one are very different and independent to guarantee a good benchmark for validation. Furthermore, the scopes of this paper are to present the AIRWAVEv2 dataset, compare its performances with respect to AIRWAVEv1 and describe the new algorithm, as described in the title.

We prefer to keep Section 5 in the Appendix to ease the readability of the paper. Following the reviewer's suggestions, we add in the revised version of the paper an estimate of the impact of additional components such as surface temperature and H2O profile.

In particular, we change the H2O profile varying randomly of 5% H2O at each level and we changed the SST randomly of 3K. When performing this test, we keep the value of the temperature of the last layer equal to the SST (no contrast). We report the results for water vapor in a new column in Table 3, while we prefer to substitute the results of the impact of varying temperature profile alone with the ones obtained when varying both SST and T profile (with no contrast between SST and temperature in the last layer). Consequently, we change the title of section 5.2 "Atmospheric temperature profile" in "Atmospheric temperature and water vapor profiles and SSTs"

In this section we rewrite the section in this way:

"One of the main error sources that can affect the AIRWAVE TCWV retrieval is the assumption of a fixed temperature profile. Actually, the atmospheric opacity is closely linked to the atmospheric density and thus to the atmospheric temperature. To estimate the impact of temperature on the retrieved TCWV, twenty different temperature profiles were randomly perturbed by ±3 K on a 1 km equi-spaced altitude grid. In order to account also for possible changes related to SST variation we change the SST accordingly to the value of the temperature in the lowest layer. Then, simulated BTs were produced with the RTM and were used to perform the TCWV retrievals for the three instruments in equatorial, mid-latitude and polar July conditions for the north hemisphere and the results were compared with respect to the unperturbed case. In the third column of Table 3 we summarise the findings of this analysis for each of the three instruments, reporting the STD of the difference both in absolute and in percentage values. The impact of these perturbations is of the order of 6% and is higher in the equatorial and mid- latitude regions and lower at the poles. Indeed, in the equatorial region, due to the higher water vapour content, the atmosphere is more opaque than at the poles so that temperature variations have a larger impact on the retrieved TCWV. These tests were also performed varying the atmospheric profile alone. Very similar but slightly higher errors are found in these cases.

Another relevant error sources can be due to differences in water vapor profile shape. To evaluate this error, we varied the water vapor profile randomly up to 5% at each atmospheric level. At maximum the impact is of 1% in equatorial case (atmospheric opacity, see the fourth column in Table 3)."

25. Table 1: What are the considered zenith angles considered here for the nadir and forward views?

They are the sub-satellite ones (0 and 55 degrees). We added this information in the Table 1 caption.

26. Table 3: Please check and correct the units. Wind speed and temperature profile cannot be in kg/m2.

In Table 3 we did not report the values of temperature and wind but their impact on the TCWV, for this reason it is in kg/m2.

27. Apart of illustrations, what are the purposes of Figs1-4? I don't see any additional validations with them? Seem they take space with a lot of redundant information. I would advise to group them in 1 single plot with 4 panels. So then you can add more validations that would be kore relevant for this paper.

These figures have no validation purpose. They show the climatology of the TCWV obtained from the AIRWAVEv2 dataset. Following the reviewer's suggestion, we grouped figures 1-4 in two figures, one with TCWV and STD maps for the four months and one for zonal and meridional means.

28. Fig5: What are the differences between "mean" (left panel) and "bias" (right panel)?

None, in the revised version of the paper we used "BIAS" in both.

29. Figs5-6-7: Please precise the time period associated with all these figures. And the considered areas for Fig.5.

These figures cover the entire (A)ATSR period from 1991 to 2012. We added this information in the revised version of the paper in figures caption.

Proposed additional bibliography

Zhao-Liang Li, Li Jia, Zhongbo Su, Zhengming Wan & Renhua Zhang (2003) A new approach for retrieving precipitable water from ATSR2 split-window channel data over land area, International Journal of Remote Sensing, 24:24, 5095-5117, DOI: 10.1080/0143116031000096014

J. A. Sobrino & M. Romaguera (2008) Water-vapour retrieval from Meteosat 8/SEVIRI observations, International Journal of Remote Sensing, 29:3, 741-754, DOI: 10.1080/01431160701311267

Lindstrot, R., Preusker, R., Diedrich, H., Doppler, L., Bennartz, R., and Fischer, J.: 1D-Var retrieval of daytime total columnar water vapour from MERIS measurements, Atmos. Meas. Tech., 5, 631-646, https://doi.org/10.5194/amt-5-631-2012, 2012.

We added Lindstrot and Li.