Author's response

Public justification (visible to the public if the article is accepted and published):

The referees have all suggested public subject to minor revisions or technical corrections, however, the number of suggested revisions is quite large. Since each of the referees has indicated a willingness to provide further review, I have decided to "reconsider after major revisions", although the authors should be aware that there is no foreseen impediment to publication at this time; the decision is only to allow the referees to provide one additional review.

We would like to thank the editor. Please, find below our responses to the Reviewers' comments and the details on how we address them in the new version of the manuscript.

The following main changes have been set with respect to the previous manuscript version, in order to answer the reviewers' comments:

Table 2 and Figure 9 have been added and the numbers of the subsequent figures and tables have changed accordingly.

The following references have been added:

Delanoë, J., and R. J. Hogan: Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. J. Geophys. Res., 115, D00H29, doi:10.1029/2009JD012346, 2010.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B.: So, how much of the Earth's surface is covered by rain gauges?. Bulletin of the American Meteorological Society, 98(1), 69-78, https://doi.org/10.1175/BAMS-D-14-00283.1, 2017.

Moreover, Figures 6, 8, 10 (11 in the revised manuscript), 11 (12 in the revised manuscript), 13 (14 in the revised manuscript), 15 (16 in the revised manuscript), and 16 (17 in the revised manuscript) have been modified.

Reviewer 1

I think the paper can be published after some clarifications about the used RT model.

We would like to thank Reviewer #1 for his/her review of our paper and the important comments and suggestions provided. Please, find below our responses to the Reviewer's comments and the details on how we address them in the new version of the manuscript.

1.1) It appears that the used model is a zeroth-order approximation of the radiative transfer equation, which can only be applied to a weakly scattering medium. The atmospheric attenuation is modeled by a one-way transmissivity parameter, which may be highly uncertain for high-frequency channels when ice and snow particles strongly scatter the upwelling emission. This class of models is applicable largely to low-frequency channels with minimal atmospheric scattering. This caveat needs to be acknowledged. Further details about the RT model seem to be necessary. The authors might consider addressing this in the appendix.

Thanks to the reviewer for the comment. We totally agree with the reviewer about the high uncertainty of the modeling of the scattering effect for high-frequency channels. However, in this work we apply the RT model in clear sky conditions, i.e. we consider only absorption and emission from atmospheric gasses and surface emission in the RT, which can be considered scatter-free for the microwave frequency range. Therefore, a comparison between the clear-sky simulated signal and the observed ones is performed, in order to highlight the snowfall signature.

In lines 103-104 (lines 93-94 in the revised manuscript) the following statement is reported:

The derived surface emissivities are used to infer the clear-sky contribution to the measured brightness temperatures (TBs) in the high-frequency channels in the snowfall retrieval process.

To make the text clearer, the term "*simulated TBs*" and the acronym "TB_{sim}" have been replaced with "*clear-sky simulated TBs*".

Reviewer 2

This work by Camplani et al. presented a new snowfall detection and intensity estimation technique. The results are very encouraging. I have several minor comments.

We would like to thank Reviewer #2 for his/her review of our paper and the important comments and suggestions provided. Please, find below our responses to the Reviewer's comments and the details on how we address them in the new version of the manuscript.

Minors:

2.1) From line 567, you explained why the newly designed method (HANDEL) performs better than SLALOM. Two reasons are provided for the better performance from HANDEL, including (1) regional database vs. global database; and (2) environmental parameters from model vs. from TBs. Which factor do you think is more important for the better performance from HANDEL?

Thanks to the reviewer for pointing out this aspect. For sure, both factors are relevant. However, we think that the added value of the approach is the use of the differences between simulated and observed TBs. In Table 3 the statistical scores evaluated for the surface snowfall rate (SSR) detection module by using different predictor sets (for the test dataset) are reported. It can be observed that the use of the differences between measured and simulated clear-sky TBs gives better performances (the fourth row of Table 3 shows an HSS=0.69) with respect to the addition of environmental variables to the predictor set (the third row of Table 3 shows an HSS=0.68). This second approach is very similar to that used in SLALOM-CT, except for the fact that HANDEL-ATMS is trained over a "regional" database, so we can state that the added value of HANDEL-ATMS derives from the use of the $\Delta TB_{obs\text{-}sim}$ (differences between measured and simulated clear-sky TB).

To clarify this point, the following statement is reported in the text (lines 448-456):

It is possible to see that the best performance is obtained when the predictor set is composed of ATMS TB_{obs} and $\Delta TB_{obs-sim}$, (besides PESCA surface flag, the pixel elevation and the cosine of the viewing angle). In particular, it is notable the improvement of the detection capabilities with respect to a predictor set composed of ATMS TB_{obs} and environmental parameters. On the other hand, the simultaneous use of both the $\Delta TB_{obs\text{-}sim}$ and the environmental parameters show scores almost equal to that obtained by using only $\Delta TB_{obs\text{-}sim.}$ This indicates that the computation of the multi-channel clear-sky TBs at the time of the overpass through the estimation of the dynamic surface class emissivity spectra and its deviation from the measured TBs plays a fundamental role in snowfall retrieval. It provides essential information to the ANN to be able to exploit the subtle snowfall-related signal in ATMS measurements. This is the most innovative aspect of HANDEL-ATMS.

which has been modified as (lines 449-459 in the revised manuscript):

*It is possible to see that the best performance is obtained when the predictor set is composed of ATMS TB_{obs} and $\Delta TB_{obs-sim}$, (besides the PESCA surface flag, the pixel surface elevation, and the cosine of the viewing angle). In particular, it is notable the improvement of the detection capabilities with respect to a predictor set composed of ATMS TB_{obs} and environmental parameters, **which is used in other approaches such as that of SLALOM-CT**. On the other hand, the simultaneous use of both the $\Delta TB_{obs\text{-}sim}$ and the environmental parameters show scores almost equal to that obtained by using only $\Delta TB_{obs\text{-}sim.}$ This indicates that the computation of the multi-channel clear-sky TBs at the time of the overpass through the estimation of the dynamic surface class emissivity spectra and its deviation from the measured TBs plays a fundamental role in snowfall retrieval, **in particular in cold/dry environmental conditions**. It provides essential information to the ANN to be able to exploit the subtle snowfall-related signal in ATMS measurements. This is the most innovative aspect of the HANDEL-ATMS.*

2.2) Figure 3 and the corresponding texts: can you explain in detail how you define the "pseudo- emissivity". Some of these emissivity values are greater than 1.

Thanks to the reviewer for the suggestion. In lines 287-288, there is the following statement:

23 GHz pseudo-emissivity (i. e. the ratio between an observed brightness temperature (TB) and a near surface temperature value) -

However, the text has been modified to make this definition clearer. In particular, lines 287-288 (lines 282-283 in the revised manuscript) have been modified

from:

23 GHz pseudo-emissivity (i. e. the ratio between an observed brightness temperature (TB) and a near surface temperature value) - pem_{23}).

to

*23 GHz pseudo-emissivity (**pem_{23}**) (i.e., the ratio between the 23 GHz observed TB and the near-surface temperature value).*

and lines 300-302 (lines 294-297 in the revised manuscript) have been modified

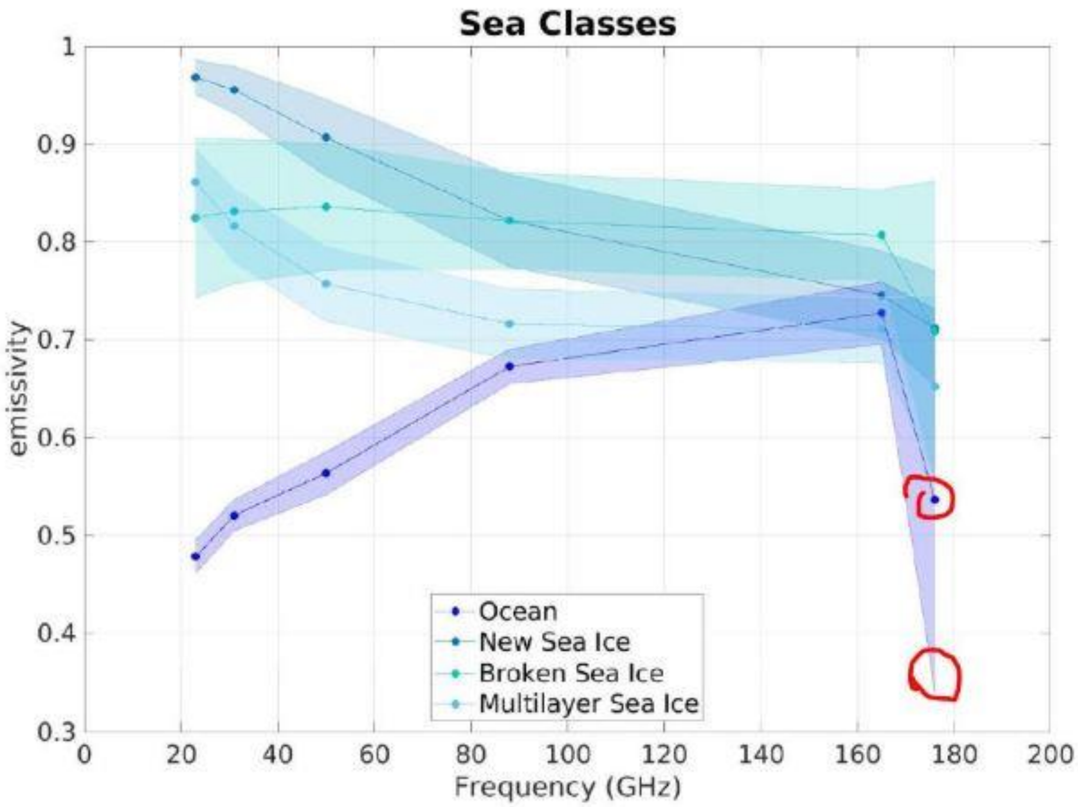
from:

Downstream of the sea ice/open water identification, information about sea ice characteristics is obtained from the analysis of the two low-frequency pseudo-emissivity (pem_{23} and pem_{31}), which are a good approximation of sea-ice emissivity for low-frequency channels especially in cold and dry conditions.

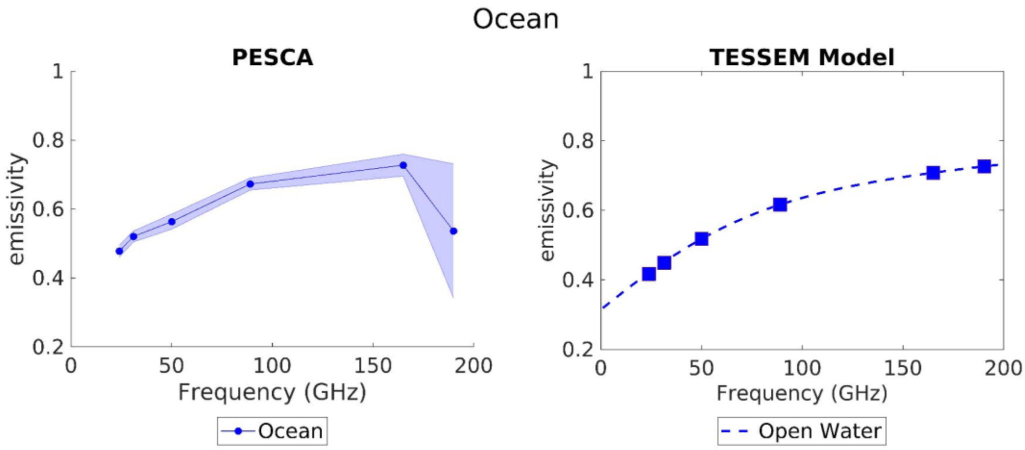
to:

Downstream of the sea ice/open water identification, information about sea ice characteristics is obtained from the analysis of the two low-frequency pseudo-emissivity values (pem_{23} and pem_{31}) (defined as the ratio between the observed TB and the near-surface temperature value) which can be considered a good approximation of sea-ice emissivity for low-frequency channels especially in cold and dry conditions.

2.3) Figure 4. It seems that the emissivity at 165 GHz is too low over ocean. See below, it can be as low as 0.35 for 165 GHz. Can you please check.



Thanks to the reviewer for the suggestion. The points highlighted in the plot correspond to 183.31 ± 7 GHz and not to 165.5 GHz. The values highlighted are the emissivity for the "Ocean" class and the lower limit of the "standard deviation" belt respectively. In the following figure, a comparison between the PESCA Ocean class emissivity retrieved values for ATMS frequency channels and the emissivity spectrum derived from the TESSEM model (*Prigent et al, 2017*) for Open Water is reported.

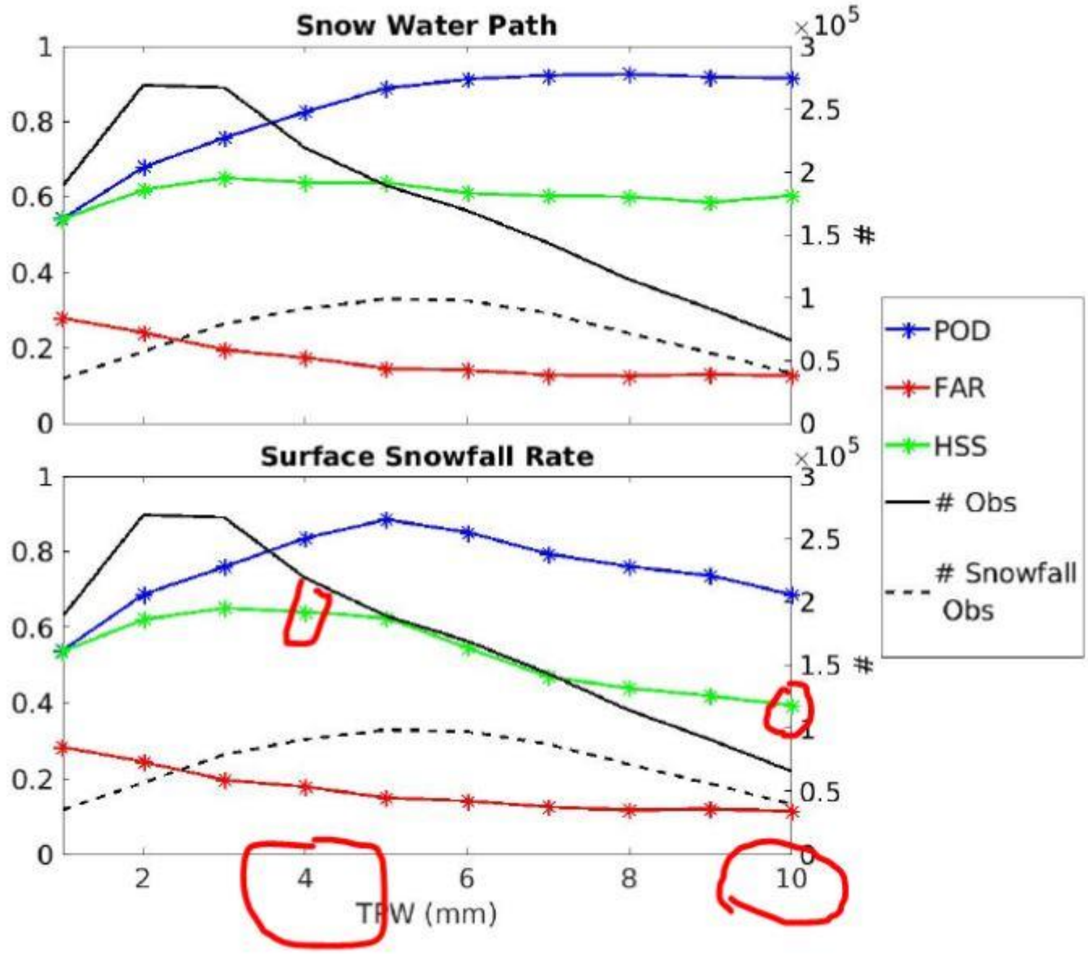


It is possible to observe that there is a good agreement up to 165.5 GHz, while at 183.31±7 GHz. On the contrary, the 183.31±7 GHz mean emissivity value is lower with respect to that obtained by applying the emissivity model, and it is characterized by high standard deviation. However, these results are preliminary, and a refinement process has been carried out by clustering each PESCA surface class. The difference between the emissivity at 183.31±7 GHz derived from PESCA and the ones in TESSEM is due to the optical thickness of the atmosphere in the water vapor absorption band. In the 183.31 GHz band, the atmosphere is opaque, due to water vapor absorption and a direct estimate of the emissivity from the observed TBs presents some issues; also downstream of the refinement process, the emissivity values obtained are lower than those expected. However, in these conditions the opacity of the atmosphere guarantees the minor impact of the surface conditions on the upwelling radiation; in fact, despite the emissivity underestimation at 183.31±7 GHz, the RMSE of the simulated clear sky TBs, as compared to the observed ones, is very small (about 3.5 K).

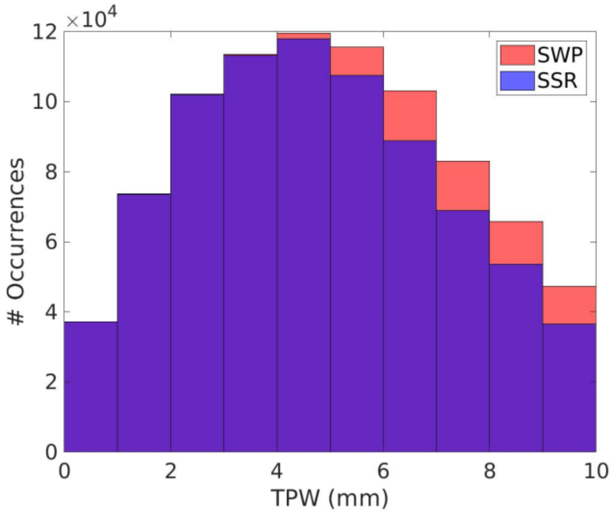
REFERENCES:

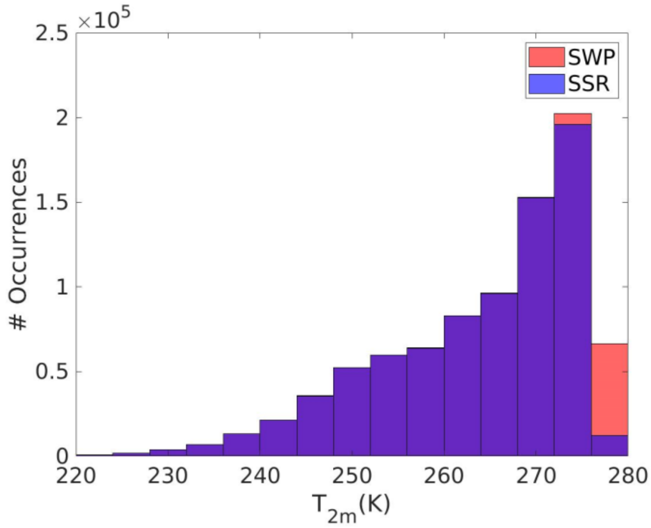
Prigent, C., Aires, F., Wang, D., Fox, S., & Harlow, C.: Sea-surface emissivity parametrization from microwaves to millimetre waves. *Quarterly Journal of the Royal Meteorological Society*, *143*(702), 596-605. https://doi.org/10.1002/qj.2953, 2017.

2.4) Figure 7 bottom panel. It shows that the HSS is smaller for TPW being 10 mm, compared with TPW being 4 mm. Specifically, the HSS decreases from about 0.6 to about 0.4. I am surprised by this result. Can you explain why? In contrast for snow water path (top panel), the HSS remains about 0.6.



Thanks to the reviewer for the question. This behavior is due to the fact that there is not a perfect correspondence between the snow water path flag and the snowfall rate flag derived from the CloudSat CPR 2C-Snow profile product, and so there are observations (about 10 % of the SWP observations in the selected datasets) where the presence of snow in the atmosphere is not matched by the presence of surface snowfall because of warmer near-surface conditions. In the following Figures the histograms of SWP/SSR occurrences as a function of TPW and T_{2m} are reported. It is possible to observe that in moister/warmer environmental conditions there is a larger number of SWP observations than SSR ones.





Generally, PMW measurements respond mostly to the snow in the atmospheric column than to snowfall at the ground, so SWP statistical scores tend to improve with increasing TPW and T_{2m} while SSR statistical scores show a maximum for TPW between 3-4 mm (or T_{2m} around 270 K) and then decrease in conditions where the mismatch between SWP and SSR become significant. In lines 473-476 (lines 477-482 in the revised manuscript) there are the following statements:

It is possible to observe that in Figure 7 SSR detection capabilities show a maximum HSS value for TPW between 3 mm and 5 mm, and then there is a slight decrease due to the decrease of POD. A similar situation can be observed in Figure 8, where HSS reaches a maximum between 250 K and 275 K, and it is lower than for SWP. ***This is due to the fact that PMW measurements respond mostly to the snow in the atmospheric column and in moister/warmer conditions the presence of snow in the atmosphere is not always linked to surface snowfall.***

In order to address the Reviewer's comment, a new Table has been added to the paper:

Class	TPW (mm)	T_{2m} (K)	# obs	% SWP obs	% SSR obs	SWP (kg m^{-2})	SSR (mm h^{-1})
Ocean	6.2	273	$3.9*10^5$	79	64	0.046	0.071
New Sea Ice	3.2	255	$2.1*10^5$	38	38	0.033	0.050
Broken Sea Ice	5.2	266	$1.4*10^5$	57	57	0.044	0.073
Multilayer Sea Ice	4.5	260	$9.9*10^4$	43	43	0.033	0.051
Land	5.3	270	$2.8*10^4$	43	41	0.043	0.068
Perennial Snow	1.6	248	$3.6*10^5$	31	31	0.022	0.035
Winter Polar Snow	2.1	245	$6.0*10^4$	32	32	0.033	0.048
Deep Dry Snow	3.8	261	$1.1*10^5$	50	50	0.040	0.066
Thin Snow	4.5	267	$1.8*10^4$	54	53	0.041	0.070
Coast	4.0	259	$3.1*10^5$	47	46	0.043	0.068

Table 2: Environmental Characteristics for each PESCA class (test dataset): the number of occurrences, the mean TPW and T_{2m} value, the percentage of SWP/SSR observations (over the total occurrences), and the mean SWP and SSR values are shown

and the following statements have been added to the text (line 342, lines 336-344 in the revised manuscript):

In Table 2 the number of PESCA class occurrences, the percentage of snowfall observations, and the most significant environmental characteristics in the ATMS-CPR coincident dataset are reported. It can be observed that Land and Ocean classes are characterized by the warmest/moistest conditions and by the most intense snowfall events (on average), while Perennial and Winter Polar Snow classes and New and Multilayer Sea Ice classes are characterized by the coldest/driest environmental conditions and by the lightest snowfall events (on average). Thin Snow and Broken Sea Ice classes show intermediate environmental conditions and snowfall intensity values. ***It is also interesting to highlight that a mismatch between the percentage of SWP and SSR observations is observed mostly over the Ocean class and, less frequently other classes (Land, Thin Snow, and Coast), where warmer and moister environmental conditions are found.***

2.5) Fig. 10, As a comparison, can you provide a similar two-panel plot from SLALOM-CT?

Thanks to the reviewer for the suggestion. These scatterplots have been already reported in the following article by the same authors of the present paper:

Sanò, P., Casella, D., Camplani, A., D'Adderio, L. P., & Panegrossi, G., A Machine Learning Snowfall Retrieval Algorithm for ATMS. Remote Sensing, 14(6), 1467, https://doi.org/10.3390/rs14061467, 2022.

Therefore, we decided not to include it in this paper.

2.6) Are these results for all ATMSs (i.e., NPP, NOAA20, and NOAA21)?

Thanks to the reviewer for the question. The study, currently, has been carried out over a dataset from 2014 to 2016, so only observations from ATMS onboard NPP were available. However, we are confident that HANDEL can be used by exploiting the ATMS measurements provided by satellites following NPP. A dedicated study is being carried out to verify if HANDEL's performance remains consistent for the other satellites.
In lines 193-196 (lines 185-189 in the revised manuscript) there is the following statement:

The present study is based on a coincidence dataset between CPR and SNPP ATMS observations between January 2014 and August 2016. The same dataset has been used for the development of SLALOM-CT (Sanò et al, 2022). Each coincidence comes from observations from CloudSat CPR and ATMS - onboard SNPP - within a maximum 15-minute time window.

However, to make this concept clearer, the text has been modified to:

The present study is based on a coincidence dataset between CPR and ATMS observations between January 2014 and August 2016. The same dataset has been used for the development of SLALOM-CT (Sanò et al, 2022). Each coincidence comes from observations from CloudSat CPR and ATMS within a maximum 15-minute time window. In the period considered within the dataset, only the SNPP satellite was in orbit, so the dataset is composed only of observations obtained from ATMS onboard this satellite.

Reviewer 3

This AMT manuscript submission describes a new ATMS Machine Learning (ML) snowfall detection algorithm (HANDEL-ATMS) that is trained on ATMS-CloudSat observations and products. This algorithm can be considered as a new retrieval scheme with strong ties to a productive lineage of microwave retrieval algorithms from this research group. The current retrieval applies an updated methodology and exploits a different sensor (ATMS) to detect and quantify snowfall rates using cross-track microwave sounder observations compared to previous related retrievals developed by this group. HANDEL-ATMS is also specifically developed to improve snowfall detection and estimation at high latitudes.
Overall, the results presented in this study are meaningful to the microwave precipitation remote sensing community and deserve to be published. The authors demonstrate that this algorithm performs well under the typically challenging conditions (light snowfall rates, very dry atmospheric conditions, surface emissivity complications) that often occur at high latitudes. Key algorithm components that enable improved algorithm performance are also described and highlighted.
I recommend that the manuscript be published after the authors consider the minor comments listed below.

We would like to thank Reviewer #3 for his/her review of our paper and the important comments and suggestions provided. Please, find below our responses to the Reviewer's comments and the details on how we address them in the new version of the manuscript.

3.1) Abstract: The first paragraph can be reduced considerably since it is covered exhaustively and effectively in the introduction. A possible way to reorganize the abstract is:

The High lAtitude sNow Detection and Estimation aLgorithm for ATMS (HANDEL-ATMS) is a new machine learning (ML)-based snowfall retrieval algorithm for Advanced Technology Microwave Sounder (ATMS) observations that is developed specifically to detect and quantify high latitude snowfall events that often form in cold, dry environments and produce light snowfall rates. ATMS and the future European MetOp-SG Microwave sounder offer good high latitude coverage and sufficient microwave channel diversity (23 to 190 GHz) that allows both surface radiometric properties to be dynamically characterized and the non-linear and sometimes subtle passive microwave response to falling snow to be detected. HANDEL-ATMS is based on a combined active-passive microwave observational dataset in the training phase, where each ATMS multichannel observation is associated with coincident (in time and space) CloudSat Cloud Profiling Radar (CPR) vertical snow profiles and surface snowfall rates. {The rest of the second abstract paragraph can follow.}

The above paragraph is only a suggestion and not mandatory. But it offers a way to distill and condense much of the introductory/background/motivation content into only 2-3 sentences.

Thanks to the reviewer for the suggestion. The text in the Abstract (Lines 8-27, lines 8-17 in the revised manuscript) has been modified as suggested
from:
Snowfall detection and quantification are challenging tasks in the Earth system science field. Ground-based instruments have limited spatial coverage and are scarce or absent at high latitudes. Therefore, the development of satellite-based snowfall retrieval methods is necessary for the global monitoring of snowfall. Passive Microwave (PMW) sensors can be exploited for snowfall quantification purposes because their measurements in the high-frequency channels (> 80 GHz) respond to snowfall microphysics. However, the highly non-linear PMW multichannel response to snowfall, the weakness of snowfall signature and the contamination by the background surface emission/scattering signal make snowfall retrieval very difficult. This phenomenon is particularly evident at high latitudes, where light snowfall events in extremely cold and dry environmental conditions are predominant. Machine Learning (ML) techniques have been demonstrated to be very suitable to handle the complex PMW multichannel relationship to snowfall. Operational microwave sounders on near-polar orbit satellites such as the

Advanced Technology Microwave Sounder (ATMS), and the European MetOp-SG Microwave Sounder in the future, offer a very good coverage at high latitudes. Moreover, their wide range of channel frequencies (from 23 GHz to 190 GHz), allows for the dynamic radiometric characterization of the surface at the time of the overpass along with the exploitation of the high-frequency channels for snowfall retrieval. The paper describes the High lAtitude sNow Detection and Estimation aLgorithm for ATMS (HANDEL-ATMS), a new machine learning-based snowfall retrieval algorithm developed specifically for high latitude environmental conditions and based on the ATMS observations.

HANDEL-ATMS is based on the use of an observational dataset in the training phase, where each ATMS multichannel observation is associated with coincident (in time and space) CloudSat Cloud Profiling Radar (CPR) vertical snow profile and surface snowfall rate.

to:

The High lAtitude sNow Detection and Estimation aLgorithm for ATMS (HANDEL-ATMS) is a new machine learning (ML)-based snowfall retrieval algorithm for Advanced Technology Microwave Sounder (ATMS) observations that is developed specifically to detect and quantify high latitude snowfall events that often form in cold, dry environments and produce light snowfall rates. ATMS and the future European MetOp-SG Microwave Sounder offer good high-latitude coverage and sufficient microwave channel diversity (23 to 190 GHz) that allows both surface radiometric properties to be dynamically characterized and the non-linear and sometimes subtle passive microwave response to falling snow to be detected. HANDEL-ATMS is based on a combined active-passive microwave observational dataset in the training phase, where each ATMS multichannel observation is associated with coincident (in time and space) CloudSat Cloud Profiling Radar (CPR) vertical snow profiles and surface snowfall rates.

3.2) Lines 42-44: I suggest offering an appropriate reference that illustrates and quantifies the lack of surface gauge coverage globally (e.g., Kidd et al. 2017).

Thanks to the reviewer for the suggestion. Lines 42-44 (lines 32-34 in the revised manuscript) have been modified from:

However, global snowfall quantification is a challenging topic in weather sciences. Ground-based instruments such as raingauges or snowgauges provide only punctual measurements which can not fully capture the spatial variability of precipitation phenomena;

to:

However, global snowfall quantification is a challenging topic in weather sciences. Ground-based instruments such as raingauges or snowgauges provide only punctual measurements which can not fully capture the spatial variability of precipitation phenomena ***(Kidd et al, 2017)****;*

and the following reference has been added to the reference section:

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B.: So, how much of the Earth's surface is covered by rain gauges?. Bulletin of the American Meteorological Society, 98(1), 69-78, https://doi.org/10.1175/BAMS-D-14-00283.1, 2017.

3.3) Lines 44-45: consider a more active writing style and shortening the sentence:"...the variability of snowflake shape and density strongly influences particle fall speed and trajectory and therefore reduces the gauge-based measurement accuracy of falling snow, especially compared to rain measurements (see Skofronick-Jackson et al 2015)"

Thanks to the reviewer for the suggestion. Lines 44-46 (lines 34-36 in the revised manuscript) have been modified from:

moreover, the variability of snowflake shape and density has a strong influence on their fall speed and trajectories and therefore gauge-based measurements of falling snow result to be less accurate than for rain (see Skofronick-Jackson et al, 2015).

to:

moreover, the variability of snowflake shape and density strongly influences particle fall speed and trajectory and therefore reduces the gauge-based measurement accuracy of falling snow, especially compared to rain measurements (Skofronick-Jackson et al, 2015).

3.4) Lines 107-110: similar to the previous comment. "Moreover, the algorithm also exploits an observational dataset composed of ATMS multichannel observations and coincident (time and space) CloudSat CPR vertical snow profiles and surface snowfall rates (hereafter the ATMS-CPR coincident dataset)". I will refrain from offering further ways to condense content and provide a more active writing style, but please know that I can provide further suggestions.

Thanks to the reviewer for the suggestion. Lines 107-110 (lines 97-99 in the revised manuscript) have been modified

from:

Moreover, the algorithm is based on the exploitation of an observational dataset where each ATMS multichannel observation is associated with coincident (in time and space) CloudSat CPR vertical snow profile and surface snowfall rate (hereafter ATMS-CPR coincidence dataset).

to:

Moreover, the algorithm also exploits an observational dataset composed of ATMS multichannel observations and coincident (time and space) CloudSat CPR vertical snow profiles and surface snowfall rates (hereafter the ATMS-CPR coincident dataset).

3.5) Line 128: It might be worth mentioning explicitly here that the CPR may struggle with high snowfall rates, but also note that CPR is uniquely suited to detect light snowfall rates that dominate high latitudes. EDIT: The authors mention high snowfall rate underestimation in Line 183, which is great. But I still think it is worthwhile to also highlight CloudSat's strength of detecting light snowfall - GPM, the only other spaceborne radar, cannot be used for training since its detection limit is far too high to effectively detect light snowfall. GPM's orbit also renders it largely useless to very high latitudes.

Thanks to the reviewer for the suggestion. See the answer to question 3.6.

3.6) Lines 189-191: This proves that I should read the entire article before commenting. My previous comment has mostly been rectified by this content. Feel free to ignore it, or at the very least explicitly highlight the further GPM drawbacks that make CPR the optimal training dataset for high latitude snowfall applications.

Thanks to the reviewer for the suggestion. The following statement has been added to the text (Line 191, lines 180-183 in the revised manuscript)

These features appear to be an advantage compared to the GPM-Core Observatory (GPM-CO), which provides observations only between 67 ° N and 67 ° S, and to the K_u- and K_a-band DPR has low sensitivity and is not suitable to effectively detect light snowfall events (Casella et al, 2017).

3.7) Line 216: The parenthetical DARDAR reference is incomplete.

Thanks to the reviewer for the comment. Lines 215-216 (lines 209-210 in the revised manuscript) have been modified

from:

The supercooled water information has been extracted from the DARDAR product (see DARDAR).

to:

The supercooled water information has been extracted from the DARDAR product (DARDAR, Delanoë &Hogan, 2010).

and the following reference has been added to the reference section:

Delanoë, J., and R. J. Hogan: Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. J. Geophys. Res., 115, D00H29, doi:10.1029/2009JD012346, 2010.

3.8) Lines 236-237: Rephrase slightly to "Moreover, clustering techniques have been used to characterize the background surface from a radiometric point of view."

Thanks to the reviewer for the suggestion. Lines 236-237 (lines 230-231 in the revised manuscript) have been modified
from:
Moreover, clustering techniques have been used to characterize from a radiometric point of view the background surface.
to:
Moreover, clustering techniques have been used to characterize the background surface from a radiometric point of view.

3.9) Lines 286-287: The inputs listed in parentheses are somewhat confusing to read due to embedded parentheses.

Thanks to the reviewer for the comment.
Lines 285-288 (lines 280-283 in the revised manuscript) have been modified
from:
*It is based on a decision tree that makes use of a limited number of inputs (the ratio TB_{23QV}/TB_{31QV} - **ratio**, the difference between TB_{23QV} and TB_{88QV} or Scattering Index - **SI**, 23 GHz pseudo-emissivity (i. e. the ratio between an observed brightness temperature (TB) and a near surface temperature value) - **pem23**).*
to:
It is based on a decision tree that makes use of a limited number of inputs: the ratio between TB_{23QV} and TB_{31QV} (ratio), the difference between TB_{23QV} and TB_{88QV} or Scattering Index (SI), 23 GHz pseudo-emissivity (pem23) (i.e., the ratio between the 23 GHz observed TB and the near-surface temperature value).

3.10) Fig. 3: I initially thought the green line indicated in the first two figure panels was somehow related to the green discriminant line indicated in Fig. 2, but I think it is the 1:1 line. Consider either explicitly mentioning this in the figure caption, or change the color or linestyle of the 1:1 line in Fig. 3.

Thanks to the reviewer for the suggestion.
The caption of Figure 3 has been modified
from:
Figure 3: Sea Ice detection and classification: relationship between 31 GHz Pseudo-Emissivity (y-axis) and 23 GHz Pseudo-Emissivity (x-axis). The color represents the mean AutoSnow sea ice percentage within each bin (top panel), the observation occurrence (middle panel), and the PESCA classification (Multi-Layer (ML), Broken and New sea ice) with the Nearest Neighbor markers (bottom panel).
to:
Figure 3: Sea Ice detection and classification: relationship between 31 GHz Pseudo-Emissivity (y-axis) and 23 GHz Pseudo-Emissivity (x-axis). The color represents the mean AutoSnow sea ice percentage within each bin (top panel), the observation occurrence (middle panel), and the PESCA classification (Multi-Layer (ML), Broken and New Sea Ice) with the Nearest Neighbor markers (bottom panel). The green continuous lines at the top and the center panels represent the bisector.

3.11) Line 339: remove the letter "e" after "constantly"

Thanks to the reviewer for the comment. Line 339 (line 333 in the revised manuscript) has been modified
from:
and constantly e throughout the year,

to:
and constantly throughout the year,

3.12) Line 394: "represents" should be regular, not superscript, font size

Thanks to the reviewer for the comment. Line 394 (line 398 in the revised manuscript) has been modified from:
where $\sigma^{\textit{represents}}$
to:
where σ represents
3.13) Line 462-463: This is somewhat of a general question, but is it worth comparing/contrasting high latitude algorithm performance versus any other ATMS snowfall (or general precipitation) retrievals that have been developed? Or other microwave retrievals? The statement provided in these lines provoked this thought. Do other precipitation retrievals provide similar statistical results (POD > 0.8, FAR < 0.2) at high latitudes?

Thanks to the reviewer for the comment. See the answer to question 3.18.

3.14) Lines 464-465 and Tables 3, 4, and 5 captions: I recommend explicitly advertising to readers what validation dataset is used to generate the statistical scores as a function of TPW and T2m. I presume CPR 2C-SNOW not used for training?

Thanks to the reviewer for the suggestion. Yes, the statistical scores have been calculated for the test dataset (see Subsection 2.3, Lines 229-234, lines 223-228 in the revised manuscript).
Lines 461-462 (lines 465-467 in the revised manuscript) have been modified (by considering that a new table - Table 2 - has been added)
from:
In Table 4 the statistical scores of HANDEL-ATMS detection module performances are reported in terms of POD, FAR and HSS
to:
In Table 5 the statistical scores of HANDEL-ATMS detection module performances are reported in terms of POD, FAR, and HSS. These statistical scores - and the plot reported in the next figures - have been calculated for the test dataset.

The caption of Figure 7 has been modified
from:
Figure 7: Dependence of HANDEL-ATMS SWP and SSR detection statistical scores on TPW. Each star represents the statistical score value for different 1-mm t bin of TPW. The left y-axis reports POD, FAR and HSS values, while the right y-axis reports the number of total and snowfall observations in the validation dataset.
to:
Figure 7: Dependence of HANDEL-ATMS SWP and SSR detection statistical scores on TPW calculated for the test dataset. Each star represents the statistical score value for different 1-mm bin of TPW. The left y-axis reports POD, FAR and HSS values, while the right y-axis reports the number of total and snowfall observations in the test dataset.

The caption of Figure 10 (Figure 11 in the revised manuscript) has been modified
from:
Figure 10: 2D Histogram reporting HANDEL-ATMS SWP (left) and SSR (right) estimation (y-axis) and 2CSP estimation (x-axis). The colorbar represents the number of observations for each HANDEL ATMS/2CSP bin. The violet dashed line represents the bisector.
to:

Figure 11: 2D Histogram reporting HANDEL-ATMS SWP (left) and SSR (right) estimation (y-axis) and 2CSP estimation (x-axis). The colorbar represents the number of observations for each HANDEL ATMS/2CSP bin (test dataset). The violet dashed line represents the bisector.

The caption of Figure 11 (Figure 12 in the revised manuscript) has been modified
from:
Figure 11: Dependence of HANDEL-ATMS SWP and SSR estimation on TPW. Each star represents the value of the statistical score for different 1-mm TPW bins. The left y-axis reports the RMSE and the mean intensity SWP and SSR value for each 1-mm TPW bin, while the right y-axis reports the relative bias, calculated as the ratio between the bias and the SWP/SSR mean value for each bin
to:
Figure 12: Dependence of HANDEL-ATMS SWP and SSR estimation on TPW calculated for the test dataset. Each star represents the value of the statistical score for different 1-mm TPW bins. The left y-axis reports the RMSE and the mean intensity SWP and SSR value for each 1-mm TPW bin, while the right y-axis reports the relative bias, calculated as the ratio between the bias and the SWP/SSR mean value for each bin

The caption of Table 3 (Table 4 in the revised manuscript) has been modified
from:
Table 3: HANDEL-ATMS SSR Detection Performance: Statistical scores for different Predictor Sets
to:
Table 4: HANDEL-ATMS SSR Detection Performance: Statistical scores for different Predictor Sets. The statistical scores have been calculated for the test dataset.

The caption of Table 4 (Table 5 in the revised manuscript) has been modified
from:
Table 4: HANDEL-ATMS detection Performance - SWP and SSR Detection Modules Statistical Scores
to:
Table 5: HANDEL-ATMS detection Performance - SWP and SSR Detection Modules Statistical Scores. The statistical scores have been calculated for the test dataset.

The caption of Table 5 (Table 6 in the revised manuscript) has been modified
from:
Table 5: HANDEL-ATMS Estimation Performance - SWP and SSR Estimation Module Error Statistics
to:
Table 6: HANDEL-ATMS Estimation Performance - SWP and SSR Estimation Module Error Statistics. The error statistics have been calculated for the test dataset.

3.15) Figs. 7, 8, and 9 and Line 467: Just to be certain that I am interpreting these figures correctly, are the POD statistics valid for the entire distribution of snowfall rates and snow water paths in each 1 mm TPW bin? It would be interesting to provide further context somewhere about how the snowfall rate and snow water path distributions vary as a function of TPW. EDIT: Fig. 11 does illustrate SWP and SSR distributions as a function of TPW. Maybe move Fig. 11 before current Figs. 7, 8, and 9 to provide more context regarding mean SWP and SSR values before the POD and FAR statistics are provided?

Thanks to the reviewer for the suggestion. Yes, your interpretation is correct. We think that moving the figure should imply some more consistent changes in the section text since, following the algorithm flowchart, the detection capabilities are analyzed first and then the estimation capabilities; however, a reference to Figure 11 (now 12) will be added.
Lines 466-467 (lines 471-473 in the revised manuscript) have been modified
from:
This is due to the combined effect of a stronger scattering signal associated with more intense snowfall events - linked to moister and warmer environmental conditions -

to:
This is due to the combined effect of a stronger scattering signal associated with more intense snowfall events - linked to moister and warmer environmental conditions, as can be observed in Figure 12 and Table 2 –

3.16) Lines 479-481: This statement is exactly what I am referring to in the previous comment. 60% POD for very light snowfall rates is excellent and should be appropriately highlighted. But I do not see how this value is derived for a 0.001 mm h-1 snowfall rate based on Figs. 7, 8, and 9.

Thanks to the reviewer for the comment. It is important to underline that detection and retrieval modules are based on different neural networks; so, the detection modules manage to identify "snowfall" conditions also in presence of very light snowfall events. In the plots below the dependence of HANDEL-ATMS snowfall detection capabilities in function of SWP/SSR values retrieved by CPR 2CSP product is reported - the statistics is calculated for snowfall observations, therefore only POD can be calculated.

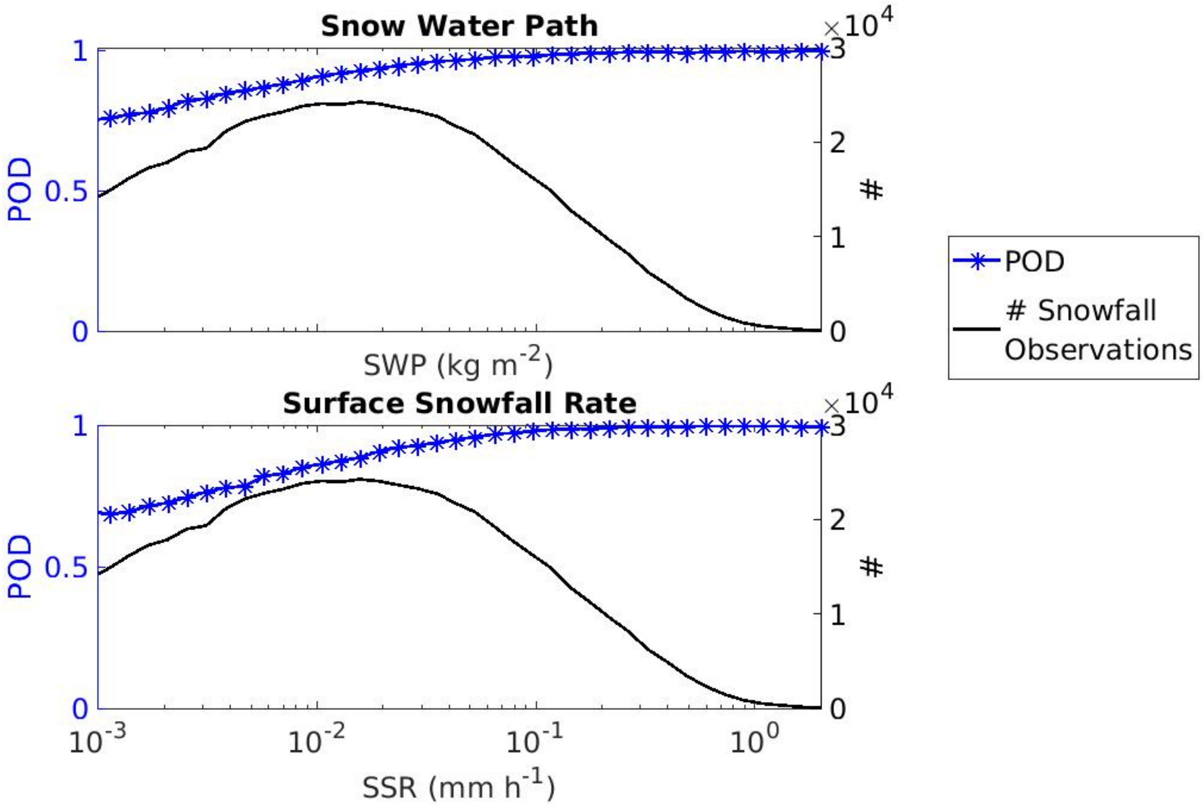


Figure 9: Dependence of HANDEL-ATMS SWP and SSR POD on SWP/SSR values. Each star represents the statistical score value for different SWP/SSR bins. The left y-axis reports POD values, while the right y-axis reports the number of snowfall observations in the test dataset. Only POD has been reported because the index has been calculated for observations where CPR 2CSP detects the presence of SWP/SSR.

This plot has been added to the paper, and Lines 481-484 (lines 485-489 in the revised manuscript) have been modified:
From:
Moreover, also considering very low SWP and SSR values (SWP ≈ 0.001 kg m^{-2}, SSR ≈ 0.001 mm h^{-1}), HANDEL-ATMS manages to detect around 60 % of the snowfall events. Similar considerations can be done also for the different background surfaces.

to:
In Figure 9 the dependence of HANDEL-ATMS snowfall detection statistical scores on SWP and SSR values retrieved by CPR 2CSP is reported. Only POD is reported because the statistics are calculated for snowfall observations only (2CSP SWP/SSR > 0 kg m-2/mm h-1). It is possible to observe that also considering very low SWP and SSR values (SWP ≈ 0.001 kg m-2, SSR ≈ 0.001 mm h-1), HANDEL-ATMS manages to detect around 60 % of the snowfall events.

3.17) Lines 487-488: Similar to the previous comment, POD > 0.7 and FAR < 0.25 for the Perennial Snow and Winter Polar Snow surface categories. These values are very impressive. But instead of generally stating that these values are impressive due to both the complicated backgrounds with variable surface emissivity and "low snowfall intensity", I recommend providing some basic quantitative guidance to bolster this analysis. A suggestion: either state what the mean or median snowfall rate is for each of these categories or provide snowfall rate distributions for various surface categories.

Thanks to the reviewer for the comment. A new table has been added to the paper in order to properly address and emphasize the important aspects raised by the reviewer.

Class	TPW (mm)	T_{2m} (K)	# obs	% SWP obs	% SSR obs	SWP (kg m^{-2})	SSR (mm h^{-1})
Ocean	6.2	273	$3.9*10^5$	79	64	0.046	0.071
New Sea Ice	3.2	255	$2.1*10^5$	38	38	0.033	0.050
Broken Sea Ice	5.2	266	$1.4*10^5$	57	57	0.044	0.073
Multilayer Sea Ice	4.5	260	$9.9*10^4$	43	43	0.033	0.051
Land	5.3	270	$2.8*10^4$	43	41	0.043	0.068
Perennial Snow	1.6	248	$3.6*10^5$	31	31	0.022	0.035
Winter Polar Snow	2.1	245	$6.0*10^4$	32	32	0.033	0.048
Deep Dry Snow	3.8	261	$1.1*10^5$	50	50	0.040	0.066
Thin Snow	4.5	267	$1.8*10^4$	54	53	0.041	0.070
Coast	4.0	259	$3.1*10^5$	47	46	0.043	0.068

Table 2: Environmental Characteristics for each PESCA class (test dataset): the number of occurrences, the mean TPW and T_{2m} value, the percentage of SWP/SSR observations (over the total occurrences) and the mean SWP and SSR values are shown

The following statement has been added (Line 342. Lines 336-344 in the revised manuscript)
In Table 2 the number of PESCA class occurrences, the percentage of snowfall observations, and the most significant environmental characteristics in the ATMS-CPR coincident dataset are reported. It can be observed that Land and Ocean classes are characterized by the warmest/moistest conditions and by the most intense snowfall events (on average), while Perennial and Winter Polar Snow classes and New and Multilayer Sea Ice classes are characterized by the coldest/driest environmental conditions and by the lightest snowfall events (on average). Thin Snow and Broken Sea Ice classes show intermediate environmental conditions and snowfall intensity values. It is also interesting to highlight that a mismatch between the percentage of SWP and SSR observations is observed mostly over the Ocean class and, less frequently other classes (Land, Thin Snow, and Coast), where warmer and moister environmental conditions are found.
Moreover, Lines 487-491 (lines 492-497 in the revised manuscript) have been modified
from:

It can be observed that, also considering specifically the classes associated to extremely dry and cold environmental conditions such as Perennial Snow or Winter Polar Snow (see Camplani et al, 2021), where the detection is more problematic due to the uncertainties in the emissivity retrieval (see Table 2) , and to the low snowfall intensity, HANDEL-ATMS has good detection capabilities (POD and FAR values greater than 0.7 and less than 0.25, respectively, for both SWP and SSR).

to:

It can be observed that, also considering specifically the classes associated with extremely dry and cold environmental conditions such as Perennial Snow or Winter Polar Snow (see Camplani et al, 2021 and Table 2), where the detection is more problematic due to low snowfall intensity (see Table 2) and to the uncertainties in the emissivity retrieval (see Table 3), HANDEL-ATMS has good detection capabilities (POD and FAR values greater than 0.7 and less than 0.25, respectively, for both SWP and SSR).

3.18) Section 4.3: This section somewhat addresses my previous suggestion of comparing other ATMS or passive microwave retrievals to the HANDEL-ATMS results. Do other passive microwave SSR retrievals exist - even historical studies - that advertise much different statistical scores than the current study? I am trying to gain further context and encourage the authors to find ways to highlight how revolutionary HANDEL-ATMS is for high latitude snowfall rate retrievals.

Thanks to the reviewer for the positive comment. An ATMS snowfall retrieval algorithm based on the CPR 2CSP product is described by *You et al, 2022*. This algorithm has been developed for snowfall retrieval over ocean, sea ice, and coastal areas and it is based on logistic regression methods. A general comparison between the two algorithms is not possible because they work over different environmental conditions (dry and cold environmental conditions typical of high latitude areas for HANDEL-ATMS, specific background surfaces for the You et al algorithm). However, it is interesting to observe that both the algorithms show higher statistical scores over open water (ocean) with respect to sea ice or a coast. Moreover, the You et al algorithm shows better performances in presence of higher SWP/SSR values. Other ATMS snowfall retrieval algorithms, such as *Kongoli et al, 2015* and *Meng et al, 2017* have been trained over a specific geographic area (the CONUS U. S.) which is not representative of the extreme high latitude environmental conditions which HANDEL-ATMS development has focused on, therefore a comparison could be not very significant. Algorithms that rely on other MW radiometers carried out by non-polar orbiting satellites, such as GMI onboard GPM-CO, do not retrieve snowfall at high latitudes, and so a direct comparison can not be carried out.

REFERENCES

Kongoli, C., Meng, H., Dong, J., & Ferraro, R.: A snowfall detection algorithm over land utilizing high-frequency passive microwave measurements—Application to ATMS. *Journal of Geophysical Research: Atmospheres, 120*(5), 1918-1932, https://doi.org/10.1002/2014JD022427, 2015.

Meng, H., J. Dong, R. Ferraro, B. Yan, L. Zhao, C. Kongoli, N.-Y. Wang, and B. Zavodsky, A 1DVAR-based snowfall rate retrieval algorithm for passive microwave radiometers, J. Geophys. Res. Atmos., 122, 6520–6540, doi:10.1002/2016JD026325, 2017.

You, Y., Meng, H., Dong, J., Fan, Y., Ferraro, R. R., Gu, G., & Wang, L.: A Snowfall Detection Algorithm for ATMS Over Ocean, Sea Ice, and Coast. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15*, 1411-1420, DOI:10.1109/JSTARS.2022.3140768, 2022