In **black** we show the original reviewer comments, in **blue** we provide the authors response

RELEVANT MODIFICATIONS

We consider that we have fully satisfied the requests of the two reviewers of the first round of reviews and also resolved the majority of concerns of the third, new, reviewer. The only remaining concern is a rather philosophical different opinion on the value of direct satellite retrievals compared to atmospheric reanalysis and integrated retrievals based on radiative transfer models. While we think that this a worthwhile debate for the scientific community it shouldn't hinder the publication of results based on well established and reliable methods accepted by large parts of the community.

The most relevant changes on the article are the following:

- Introduction of a comparison of AMSU-B data to ERA5 reanalysis in new Section 3.3 and a comparison of ERA5 to GPS and radiosondes at the end of Section 3.2. With this, we show that the performance of our retrieval shows similar quality as the ERA5 reanalysis and hence reinforce its usefulness
- Expansion of the Sections 2.7 and 4 regarding the masking of ice cloud artefacts. In Section 2.7 we provide a more detailed description of the method and justification of its need. In Section 4, we provide some additional statistics and expanded/developed further the analysis
- The other changes in the article deal with the minor comments from both reviewers, fixing typos and rephrasing or expanding the relevant sentences (such as the acknowledgements)

COMMENTS REVIEWER 1

1) Please acknowledge provision of radiosonde and GPS data and provide information to where you got the data from (e.g., in the acknowledgment).
P12 L17-20: Added the following acknowledgment: The authors would like to thank the Department of Atmospheric Science, University of Wyoming for the radiosonde raw data (http://weather.uwyo.edu/upperair/sounding.html) and the International GNSS Service for the GPS data (ftp://igs.ensg.ign.fr/pub/igs/data/).

2) page 2, line 31: ASMU->AMSU
P2 L28: Corrected

3) p2, l32: I think it should say RMSD and not standard deviation. If so, please change.
P2 L28: Yes, it should. Corrected

4) p3, l11: Eumetsat->EUMETSAT
P3 L6: Corrected

5) p5, l23: wich->which
P5 L23: Corrected

6) p8, l17: Please add „associated with the quasi periodic peaks“ after outliers.
P8 L15: Added

7) p8, l22: It is not clear to me what is meant with „homogenized year“. Please rephrase.
P8 L21: the word 'year' after homogenized is a typo from previous rephrasing, it has been removed

8) p8, l34: I think this statement is maybe not correct. The bias and the RMSD are small in winter because the values are small. The reason for a low correlation is likely that the temporal coherence is less pronounced.
P9 L9-12: We agree with your statement but do not see how it is in disagreement with what we wrote about that the higher summer RMSD is also seen in Fig. 8. We have reformulated the sentence and added your statement.

9) p9, l6,7: Please remove „The bias…datasets.“
P9 L7: Removed

COMMENTS REVIEWER 2

The authors present an update to a water vapour retrieval suitable for the polar regions, expanded to a more modern sensor (MHS) and now including a screening procedure intended to mitigate the deleterious effects of significant scattering from clouds on the retrieval. The paper presents itself as describing these two 'advances' to the previous retrieval, 'intended as groundwork' for a future planned combined product that would incorporate oceanic microwave imager retrievals as well, combining the two into a pan-Arctic product that could potentially cover a long time period. While the work is nicely presented and well written, I do not believe that this rises to the level of significance to the community that merits publication in this journal. For that reason, I recommend rejection of the manuscript as it stands, with encouraged resubmission if the authors follow through on the stated future work. More in-depth comments follow, broken up into a few major bullet points and minor comments.

1. The abstract lays out the paper as presenting two advances to an old and established retrieval of Arctic water vapour. The first is simply expanding the old retrieval to a new sensor, which is almost a trivial exercise since the channels are almost identical and the difference of absorption characteristics between the 150 vs. 157GHz channels is fairly trivial. It is nice to see that the retrieval works similarly for MHS as it does AMSU-B, and this is well laid out by the authors, but it is not surprising or noteworthy for the community that reads AMT -- perhaps a small technical challenge but not scientifically significant.

P1 L1-10, Section 3: We agree that it is not surprising that the adaptation of the retrieval to MHS works. As you, however, correctly describe there are differences between the two sensors like the different frequencies and polarization in some channels. We here not only describe the technical part of the adaptation but also prove its success with independent data comparisons. We consider both parts, the adaptation and evaluation, scientifically significant and think that they need to be published so it provides a reference for the extended time series.

The abstract's second advance touted is a new screening for artefacts caused by convective clouds. While this is potentially quite interesting, it is essentially a footnote in Section 2 of the manuscript, and the description of this new filtering method is literally restricted to 5 lines of the total manuscript (P6 L23-27). Furthermore, I am not convinced that it is necessarily filtering out 'high cloud ice content in convective clouds' as the abstract states; rather the authors infer that the low retrieved TWV is indicative of a scattering signal from cloud ice, but this is not demonstrated in the paper and thus it appears that it is just an assumption.It could be justified as such if compared to other satellite imagery or an IWP product.

There is clear evidence that there is an effect of ice clouds at the AMSU-B frequencies as studied in Sreerekha (2005), and has been used related to the low retrieved TWV values in different articles such as (Melsheimer et al., 2016). We have added some additional explanations to justify this need better:

P6 L16-32, Section 2.7: We have expanded the justification of the method

P10 L2-14 Section 4: We have added ERA5 maps to Figure 11 to show visually the problematic regions, i.e., ice clouds, in AMSU-B that are not present in ERA5

The remainder of the paper holds no strong conclusions: 'The improved retrieval performs better when compared to another satellite product and to in situ data' (P9 L32) and 'the results are satisfactory' (P10 L5). It is unclear what exactly is demonstrably better as much of the discussion is qualitative, or even that the updated retrieval could outperform reanalysis datasets, which is an almost necessary test for retrievals to demonstrate.

P9 L13-20, Figure 8: Assessed performance of ERA5 compared to radiosondes and GPS

P9 L21-32, Figure 10 and Table 5: Added new section 3.3, with a comparison with ERA5

2. The second major comment has more to do with the methodology upon which the study rests. This is a subjective opinion, but so-called 'non-physical' retrievals such as Miao et al. (2001) were quite important twenty years ago when radiative transfer codes were slower and less advanced, but are becoming less useful today. The results show the downsides to such a regression-based bin method, with big gaps visible in Fig. 1. Why would the community use such a product when reanalyses have no such gaps in coverage or artefacts between bins, not to mention blended TWV products that exist

too? Modeling of sea ice emissivity is of course still a big challenge, but physically-based retrievals from microwave radiometers already exist over sea ice and indeed all surfaces (for example, see the MIRS retrieval from NOAA: https://www.star.nesdis.noaa.gov/mirs/geonwp.php). If regression-based retrievals such as the one presented are to remain relevant, they need to demonstrate their worth relative to similar products, including reanalyses (see e.g. Duncan and Kummerow 2016). If this paper had presented the validation against in-situ sources alongside comparison with say ERA5 data and shown that it outperforms the reanalysis, then it is of much more interest to the community. Even the proposed combined TWV product of this retrieval with RSS data (P10 L13) would need to prove this, and it is as of yet far from certain; the bin-based artefacts are a major concern and merging with RSS would be difficult in itself due to their own biases and simplifying assumptions made for radiative transfer. If the methodology can be shown to outperform physically-based retrievals (with a full forward model) then it has interest for the community, but otherwise it strikes me as requiring corrections on top of corrections that do not lead to greater physical understanding, and could be perhaps be better accomplished by a neural net retrieval.

We believe that direct retrievals like the one presented here are still relevant for the scientific community and operational centers, which is also demonstrated by the use of such data in recent publications (using AMSR-E TWV over ocean or the AMSU-B TWV retrieval that we deal with in this paper summarized below). As the reviewer mentions this point is more a "subjective opinion" and should not dominate the assessment of this manuscript. Our retrieval is based on physical principles, which due to the complexity of the sea ice - atmosphere microwave emission need simplifications based on empirical factors. While RTM based retrievals are more complex they also use approximations and simplifications. A descrimination in physical and nonphysical retrievals is not correct. The reviewer already mentions one advantage of such direct retrievals: they are fast. Another advantage is that they do not need any auxiliary data. Thus they are independent and completely based on the satellite measurements.

Recent publications using AMSR derived TWV (http://www.remss.com/missions/amsr/):

Casadio, S.; Castelli, E.; Papandrea, E.; et al. (2016): Total column water vapour from along track scanning radiometer series using thermal infrared dual view ocean cloud free measurements: The advanced infra-red water vapour estimator (airwave) algorithm. Remote Sensing of Environment, 172, 1-14. doi:10.1016/j.rse.2015.10.037.

Grossi, M.; Valks, P.; Loyola, D.; et al. (2015): Total column water vapour measurements from gome-2 metop-a and metop-b. Atmospheric Measurement Techniques, 8, 1111-1133. doi:10.5194/amt-8-1111-2015.

Recent publications using AMU-B derived TWV:

S. A. Buehler, S. Östman, C. Melsheimer, G. Holl, S. Eliasson, V. O. John, T. Blumenstock, F. Hase, G. Elgered, U. Raffalski, T. Nasuno, M. Satoh, M. Milz, and J. Mendrok. A multi-instrument comparison of integrated water vapour measurements at a high latitude site. Atmospheric Chemistry and Physics, 12(22):10925–10943, 2012. (doi:10.5194/acp-12-10925-2012)

M. Palm, C. Melsheimer, S. Noël, J. Notholt, J. Burrows, and O. Schrems. Integrated water vapor above Ny Ålesund, Spitsbergen: a multisensor intercomparison. Atmospheric Chemistry and Physics, 10(3):1215–1226, 2010.

A. Rinke, C. Melsheimer, K. Dethloff, and G. Heygster. Arctic total water vapor: Comparison of regional climate simulations with observations and simulated decadal trends. Journal of Hydrometeorology, 10(1):113–129, 2009. (doi:10.1175/2008JHM970.1)

We agree with the assessment that such direct retrievals need to show that they are useful and can compete with model based datasets like the ERA5 reanalysis. Therefore we extended the evaluation of our TWV dataset by a comparison with ERA5:

P9 L13-20, Figure 8: Assessed performance of ERA5 compared to radiosondes and GPS

P9 L21-32, Figure 10 and Table 5: Added new section 3.3, with a comparison with ERA5

However, the ERA5 assimilates all radiosonde data and thus no independent evaluation of ERA5 is possible. It is therefore not surprising that ERA5 and the radiosonde agree well. As the GPS TWV measurements are taken at the same locations as the radiosonde launches, also the agreement to GPS is good. While our direct retrieval does not outperform the ERA5 dataset it does show similar quality. As for our dataset the evaluation is completely independent from the in-situ data and thus meaningful, we can assume that a similar performance is also achieved at other geographical locations. This conclusion cannot be drawn for the ERA5 dataset because of its inherent dependence on the radiosondes measurements. This actually demonstrates another advantage of direct, independent retrievals like the one presented here.

3. The radiative transfer equation upon which the methodology rests struck me as maybe being incorrect (Eq. 1). If we take the case of surface emissivity of 1, then TB is directly proportional to surface temperature; if we take a fully opaque atmosphere with negligible transmittance (tau>>1), then again the second term goes to zero and TB is again directly proportional to Ts; if surface emissivity were zero, then TB is essentially Ts minus an atmospheric contribution? I apologise if I am misinterpreting this, but it makes no sense to me when I consider these cases. However, it is indeed the exact same equation given in Miao et al. (2001) and originally in Guissard and Sobieski (1994), so I am perplexed. I did not have the time to follow the full derivation in the G&S 1994 paper, but it seems suspect to me. I would suggest examining this in detail to make sure this isn't a typo, because it appears like a form given in Grody (1976) but with Ts and To flipped. Again, apologies if I have misinterpreted this--it just struck me as odd.
P3 L21-23, Equation 1: At the end of the authors comment (section of this document titled "Radiative Transfer Equation") a detailed explanation is given about how the equation is correct. We have slightly expanded the description of the term mp to make it less ambiguous. It now reads: "..., and m_p a correction to take into account both a non-isothermal atmosphere and the difference between the surface (skin) temperature, T_s, and the temperature of the atmosphere at the ground, T_0, ($m_p = 1$ would be the isothermal case and $T_0 = T_s$)""

Minor comments:
P1 L12: The title uses 'polar' but the paper almost exclusively uses 'Arctic' only. Unless there is some focus on the Southern Hemisphere too the title should be reconsidered.
P1L12 Changed it to Arctic

P2 L1: Is 1m squared a typo?
P1L21 No, it is not. Changed the definition to the phrasing "when mentioning atmospheric water content, we refer to the vertically integrated mass in an air column with an area of 1 m^2" to make it more obvious

P2 L10: Fix citation Bobylev and Mitnik
P2L6 Corrected

P2 L16: According to OSCAR SSM/T2 confusingly stands for Special Sensor Microwave Humidity (https://www.wmo-sat.info/oscar/instruments/view/535)
P2L12 Changed it

P2 L20-21: Is there proof of this statement? A citation or elaboration would be good here.
P2 L16-18 The problematic statement is "Above this value, two of the 183.31 GHz band channels become saturated and the sensor is not able to "see" through the whole atmospheric column anymore". We cited the relevant work (see below) and added a short sentence elaborating on this: "In other words, when the TWV reaches a certain threshold, the brightness temperature at these AMSU-B channels does not change with increasing TWV".
The relevant citations are:

- *J. Miao, "Retrieval of atmospheric water vapor content in polar regions using spaceborne microwave radiometry," Alfred-Wegener Inst. Polar Marine Res., Bremerhaven, Germany, 1998.*
- *Melsheimer, C. and Heygster, G.: Improved Retrieval of Total Water Vapor Over Polar Regions from AMSU-B Microwave Radiometer Data, IEEE Trans. Geosci. Rem. Sens., 46, 2307–2322, 2008.* Specifically the following paragraph as a more extended explanation: "we use the term "saturation" in the following sense. When we measure a quantity T in order to determine a parameter W, this of course implies that T does really depend on W, i.e., T = T(W). In reality, this

often applies only to a limited range of W. Typically, when W reaches a certain threshold, T does not change with increasing W anymore. This is generally called saturation. In our case of T being the brightness temperature at one of the AMSU-B channels and W being TWV, T first increases with W, then levels off and starts to decrease slowly (Miao, 1998). This means that at low TWV, the brightness temperature is dominated by the thermal emission of the water vapor and, hence, increases with the water vapor content. Beyond a certain TWV value, the atmosphere becomes opaque, and the brightness temperature comes from the upper part of the atmosphere. The higher the TWV content, the higher—and colder—the portion of the atmosphere that contributes to the brightness temperature; hence, the brightness temperature decreases with increasing water vapor. Beyond a certain TWV value, the atmosphere becomes opaque, and the brightness temperature comes from the upper part of the atmosphere. The higher the TWV content, the higher—and colder—the portion of the atmosphere that contributes to the brightness temperature; hence, the brightness temperature decreases with increasing water vapor content."

P3 L10: Is a table with launch dates necessary? It does not really impact the paper.
P3L8 This was a modification introduced during the last round of revision on the suggestion of a reviewer. Although this table is not required to understand the paper, we feel it provides a more clear explanation about the different coverage periods of the different sensors and satellites. While these dates are available online, it is not straightforward to find them together.

P3 L16: Typo in citation, Sobieski
P3L17 Corrected

P4 L15: What are the units on k? Since absorption coefficients for water vapour are very well known, the derived regression parameters C could be compared against values in the literature
P4L5 Added units on k (m^2/kg). Thus, the product of mass absorption coefficient and total water vapour (kg/m^2), which is the optical depth, is dimensionless as it should be.. We agree that absorption coefficients are well known. However, the relation between the mass absorption coefficients and the the regression coefficients is complicated and has never been formulated explicitly until now (for each sub-algorithm, the two regression coefficients depend on the three mass absorption coefficients of water vapour at the three frequencies used, see details in *Miao, 1998*, cited above). Thus, even a check if the regression coefficients are consistent with water vapour mass absorption coefficients is beyond the scope and aim of this study.

P5 L17: Perhaps I missed this, but does the manuscript state how the 'surface types are obtained'? This is a key part of the algorithm and surely any future combined product. There is something at P7 L25, but it is unclear if this is how the algorithm functions or if that was just for that particular analysis.
P3 L13-15 Changed the statements in P7 L25 to start of Section 2.1 (P3 L13-15) to reflect that this is how the algorithm functions

Section 3.1: How are coincident points defined?
P7 L13-14, Section 3.1: Now the text explicitly mentions this: "For this analysis, we considered all the coincident points in the daily gridded data with a 0.25º grid."

P7 L9: Is there any justification for saying that time differences are 'likely' the cause of differences, or is this speculation?
P7 L17-19: During the initial analysis, we looked at different parameters that could be the source of these TWV differences, such as the time difference, possible gridding issues and the sub-algorithm associated with each point. We found out that a significant majority of the "high TWV difference" points are associated with time differences. We checked this again, and came to the same conclusion. The word "likely" is probably wrongly used here, since there is an actual justification for this. We changed this sentence in the paper to: "These points are mostly associated with time differences of the satellite overpasses, and amount to only about 0.27% of the data, so they are not significant in the overall picture."

P7 L12: What was the 'expected amount of data'? I found this confusing.
P7 L21: Removed this since it is confusing and not relevant here

P7 L19 It would be interesting to investigate why there is this ‘low agreement in summer’ rather than just to ‘presume’ -- this could possibly be tested by contrasting open water with retrievals over ice.
P7 L25-28: We do something along the lines of this suggestion in the “surface study” (Figures 2 and 3) We agree that it would be interesting to go even more in-depth with this, but though we indicate the agreement is lowest it is still very good and it feels unneeded

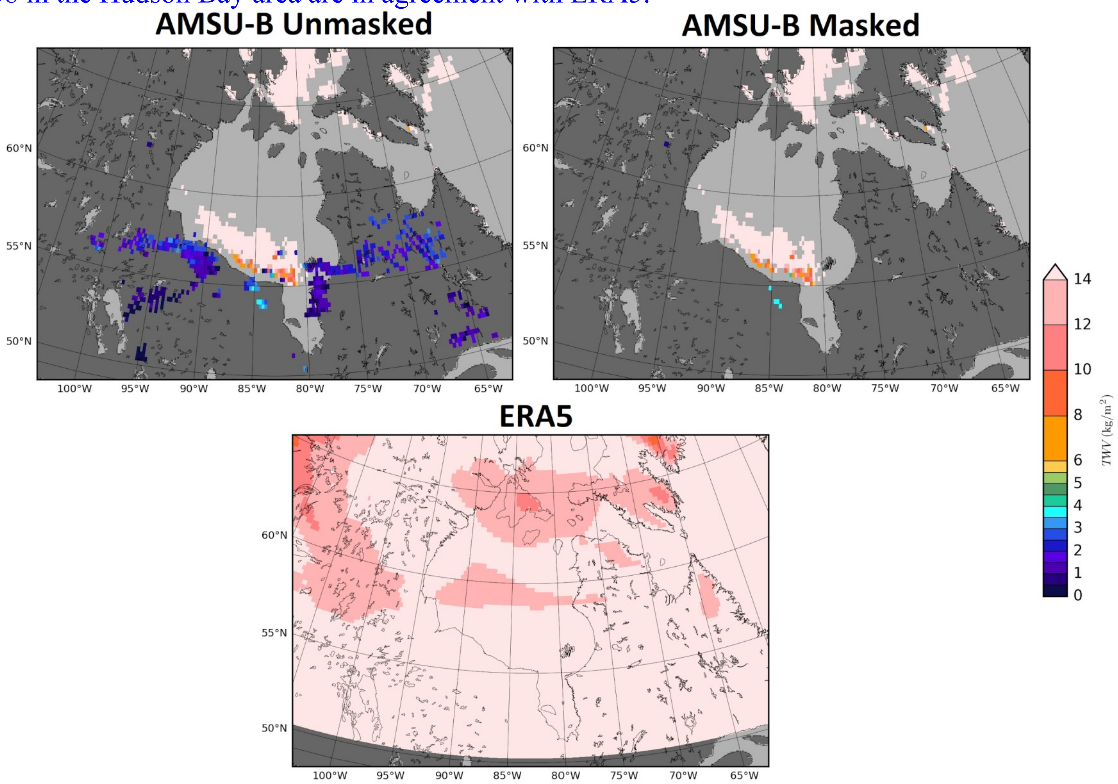
P8 L5: I don’t understand this -- you eliminated the outliers from the analysis and then found that there was good agreement? What was the justification for eliminating the outliers?
P8 L13-14 Rephrased, added elaboration why these outliers should be eliminated and are problematic

P8 L30: The bias values should be smaller than RMSD by definition.
P9 L5-6 Removed sentence.

P13 L5: Typo ‘Anctarctica’
P14 L15 Corrected

Fig. 10: I really like the colour scale used, but it seems insufficient for the July panels.
Suggest using separate colour scales, one for each season so that patterns over sea ice can be seen in both seasons.
Fig 10: We consider this a good suggestion, and will implement it for the following versions of the paper.

Fig. 12: Some discussion of the third row here seems necessary. Surely it’s not physical to expect TWV=14 or more in the southern Hudson bay with TWV<3 just to the south even after screening?
Fig 12: If we compare ERA5 reanalysis with the AMSUB TWV on that date (you can see the figure below for the date and region in question, 6.08.2008), we can see that the expected - physical and correct - TWV is the one higher or equal to 14kg/m^2, both retrievals are in agreement about these values. The remaining couple of pixels in the AMSUB retrieval with ~3.5kg/m^2 are just not screened by the mask and are unphysical. We added a short paragraph summarizing this in the analysis of Figure 12, P L “We confirmed by comparison to the ERA5 atmospheric reanalysis that the remaining high TWV values are within the expected range. Also the high, >14 kg/m^2, TWV values on 6th July 2008 in the Hudson Bay area are in agreement with ERA5.”



Radiative Transfer Equation

The reviewer's comment, item 3., seems to show that the brightness temperature seen by the satellite is just proportional to the ground temperature for a ground with emissivity equal to 1, but also for a totally opaque atmosphere (transmittance equal to 0), which would obviously be wrong. The error in the reviewer's argumentation was to assume that the factor m_p would be the same in both cases. Here is a detailed derivation:

The equation in question, i.e., equation (1) in our manuscript, corresponds to equation (27) in Guissard and Sobieski (1994): It describes the brightness temperature at the satellite, above the atmosphere (at $z = H$):

$$T_B(\theta) = m_p T_s - (T_0 - T_c)(1 - \varepsilon_s)\mathrm{e}^{-2\tau} \tag{1}$$

where

- τ is the total optical depth of the atmosphere, i.e. the integral of the atmospheric extinction profile $\alpha(z)$: $\tau = \int_0^H \alpha(z)\mathrm{d}z/\cos\theta$; note that Guissard and Sobieski (1994) use $\Upsilon = \mathrm{e}^{-\tau}$ instead
- T_0 is the atmospheric temperature at the lowest level, $T_0 = T_a(z=0)$ (Guissard and Sobieski (1994) call it T_a)
- T_s is the surface (skin) temperature, in general not equal to T_0
- T_c is the 3K cosmic background radiation

The whole complexity of radiative transfer is in the factor m_p, defined in Guissard and Sobieski (1994), equation (28). It contains two effects: (1) the deviation of the atmosphere from being isothermal, and (2) the fact that $T_s \neq T_0$. The definition is

$$m_p = 1 + \left[(1 - \varepsilon_s)\mathrm{e}^{-\tau}\frac{T_0 - T_s}{T_s} - \frac{I_p}{T_s}\right] \tag{2}$$

where

$$I_p = I_1 + (1 - \varepsilon_s)\mathrm{e}^{-2\tau} I_2 \tag{3}$$

and the terms I_1 and I_2, defined in Guissard and Sobieski (1994), equations (21) and (25), both are vertical integrals related to the atmospheric absorption, weighted by the vertical gradient $T_a'(z)$ of the atmospheric temperature $T_a(z)$, and both vanish for an isothermal atmosphere. It is important to note that $m_p = 1$ holds only if the atmosphere is isothermal *and* $T_0 = T_s$.

Now we set the ground emissivity ε_s equal to one, so the brightness temperature at the satellite is just

$$T_B(\theta) = m_p T_s \tag{4}$$

However m_p is not equal to one, but still contains several terms related to atmospheric emission and absorption:

$$m_p = 1 + (1 - \mathrm{e}^{-\tau})\frac{T_0 - T_s}{T_s} - \frac{I_1}{T_s} \tag{5}$$

Note that I_1 vanishes for an isothermal atmosphere. Inserting this into our equation (4) above, we in fact get, for the brightness temperature at the satellite, for $\varepsilon_s = 1$ and non-opaque atmosphere:

$$T_B(\theta) = m_p T_s = T_s \mathrm{e}^{-\tau} + T_0(1 - \mathrm{e}^{-\tau}) + I_1 \tag{6}$$

The first term describes the emission by the ground, transmitted through the atmosphere, the second term is the upwelling atmospheric radiation for an isothermal atmosphere at temperature T_0 and the third term is a correction of the upwelling radiation for a non-isothermal atmosphere.

If, in contrast, we just set $\tau \gg 1$, so we have an opaque atmosphere and $\mathrm{e}^{-\tau} \approx 0$, we also get

$$T_B(\theta) = m_p T_s \tag{7}$$

as above, but the difference lies in m_p:

In this case, it is

$$m_p = 1 + \frac{T_0 - T_s}{T_s} - \frac{I_1}{T_s} \tag{8}$$

So we get, for the brightness temperature at the satellite

$$T_B(\theta) = m_p T_s = T_s + (T_0 - T_s) + I_1 = T_0 + I_1 \tag{9}$$

Here, T_0 is the upward emission of a totally opaque atmospheric layer of uniform temperature T_0, and I_1 is then the correction term for the temperature not being uniform.

This should resolve item 3. We have noticed that our brief description of the factor m_p after equation (1) in our manuscript was not complete and have corrected that. We now state:

"..., and m_p a correction to take into account a non-isothermal atmosphere and the difference between the surface (skin) temperature, T_s, and the temperature of the atmosphere at the ground, T_0, ($m_p = 1$ would be the isothermal case and $T_0 = T_s$)"

Improved Water Vapour retrieval from AMSU-B/MHS in ~~polar regions~~the Arctic

Arantxa M. Triana-Gómez[1], Georg Heygster[1], Christian Melsheimer[1], Gunnar Spreen[1], Monia Negusini[2], and Boyan H. Petkov[3]

[1]Institute of Environmental Physics, University of Bremen, Bremen, Germany
[2]Institute of Radio Astronomy, INAF, Bologna, Italy
[3]Institute of Atmospheric Sciences and Climate, CNR, Bologna, Italy

Correspondence: Arantxa M. Triana-Gómez (aratri@uni-bremen.de)

Abstract.

~~Exact monitoring~~ Monitoring of water vapour in the Arctic on long time scales is essential for predicting Arctic weather and understanding climate trends, as well as addressing its influence in the positive feedback loop contributing to Arctic Amplification. However, this is challenged by the sparseness of in-situ measurements and the problems that standard remote-sensing retrieval methods for water vapour have in Arctic conditions. Here, we present advances in a retrieval algorithm for vertically integrated water vapour (total water vapour, TWV) in polar regions from data of satellite-based microwave humidity sounders: (1) In addition to AMSU-B (Advanced Microwave Sounding Unit-B), we can now also use data from the successor instrument MHS (Microwave Humidity Sounder); (2) artefacts caused by high cloud ice content in convective clouds are filtered out. Comparison to in-situ measurements using GPS and radiosondes during 2008 and 2009 as well as to radiosondes during the N-ICE2015 campaign and to ERA5 reanalysis show overall good performance of the updated algorithm. ~~Combining TWV data from the present algorithm with those retrieved from microwave imagers like AMSR-E and AMSR2 makes a continuous record of TWV since the year 2000 possible, with nearly complete and year-round coverage of the Arctic.~~

1 Introduction

Water vapour is a key element of the hydrological cycle (Chahine, 1992; Serreze et al., 2006; Jones et al., 2007; Hanesiak et al., 2010), with shifts in it affecting atmospheric transport processes, creating and intensifying droughts and flooding (Trenberth et al., 2013). Additionally, as the most important greenhouse gas in the atmosphere, it has a dominant effect on climate and radiative forcing (Soden et al., 2002; Dessler et al., 2008; Kiehl and Trenberth, 1997; Trenberth et al., 2007; Ruckstuhl et al., 2007). Hence, it is essential to monitor its variability considering both that water vapour increases when temperature does and the anthropogenic increase of other greenhouse gases (Solomon et al., 2010), with the water vapour positive feedback loop highlighted as part of other feedbacks responsible for Arctic Amplification (Francis and Hunter, 2007; Miller et al., 2007; Screen and Simmonds, 2010; Ghatak and Miller, 2013). In summary, understanding the water vapour cycle has high value, yet our comprehension is incomplete (Stevens and Bony, 2013). Throughout this paper, when mentioning atmospheric water content, we refer to the vertically integrated mass in ~~a column of air with a base~~ an air column with an area of 1 m^2, and call it total water vapour (TWV, sometimes also called column water vapour, integrated water vapour or total precipitable water), the units are hence kg/m^2.

Balloon-borne radiosondes are a standard method for retrieving the water vapour profile. Additionally, ground-based retrievals by microwave radiometers as well as GPS-based retrievals – while having a lower vertical resolution – are good for monitoring purposes in regions where ground stations can be installed. However, in the Arctic, neither radiosondes measurements nor ground-based retrievals are sufficient for this purpose because weather stations are too scarce. Only satellite measurements fulfill the global coverage requirements. An additional challenge is to construct a consistent long-term climate record, due to the changes in measuring instruments, and degradation of the existing ones. Because of the strong absorption properties of water vapour in the infrared and microwave range, suitable space-borne instruments can in principle ensure a complete global coverage of water vapour retrievals (Miao et al., 2001; Bobylev et al., 2010). In polar regions, however, satellite retrieval of water vapour faces a number of obstacles such as cloud cover which restricts infrared measurements, or incomplete understanding of the high and highly variable sea-ice emissivity which challenges microwave measurements. Some studies – like the one by Weaver et al. (2017) – have been done for TWV in the Arctic atmosphere, but none of them have been able to provide a long-term Arctic-wide data set.

An important step for Arctic water vapour retrieval comes from the work of Miao et al. (2001). They used data from the SSM/T2 (Special Sensor Microwave /Temperature 2Humidity) humidity sounder to develop an algorithm which was designed to work in the Antarctic. The key concept of this method is the use of several microwave channels with similar surface emissivity but different water vapour absorption. These are the three channels near the 183.31 GHz water absorption line (183.31 $\pm$1, $\pm$3 and $\pm$7 GHz), which, together with the channel at the 150 GHz window frequency, allows retrieval of TWV values up to about 7 kg/m^2. Above this value, two of the 183.31 GHz band channels become saturated and the sensor is not able to "see" through the whole atmospheric column anymore. In other words, when the TWV reaches a certain threshold, the brightness temperature at these AMSU-B channels does not change with increasing TWV (?Melsheimer and Heygster, 2008). This limited range is enough for Antarctica, and suffices for the Arctic in winter conditions (in the polar winter atmosphere, the water vapour column is typically around 3 kg/m^2 according to Serreze et al. (1995)), as well as for the central Arctic (above 70° N) most of the year. However, because of the upper limit, this method cannot ensure monitoring of the complete yearly cycle. The algorithm developed by Melsheimer and Heygster (2008) extends the TWV retrieval range over sea ice by including the AMSU-B (Advanced Microwave Sounding Unit-B) 89 GHz channel into the retrieval. Using the triplet of the 183.31$\pm$7, 150 and the 89 GHz channels allows the retrieval to function up the saturation limit of the 183.31$\pm$7 GHz channel. This method has been compared with other datasets: In Rinke et al. (2009) a comparison with the HIRHAM model showed realistic patterns and maximum root-mean-square differences for monthly data in summer of 1-2.5 kg/m^2. For the comparison with Ny Ålesund radiosondes in Palm et al. (2010), the correlation coefficient was 0.86 and the slope 0.8$\pm$0.04 kg/m^2. And lastly, in Buehler et al. (2012) ASMU-B AMSU-B TWV are compared to GPS data from Kiruna, with standard deviations RMSD of 1 kg/m^2 and a correlation coefficient of 0.86. However, the AMSU-B algorithm is not without problem: while the frequency range allows it to bypass most clouds, the AMSU-B sensor is still sensitive to convective clouds with high ice content. Here we provide an approach for filtering out problematic data caused by the effect of such ice clouds. This is intended as groundwork for the planned merging with TWV retrieved over open ocean based on passive microwave imagers (product described by Wentz and Meissner, 2006).

In Section 2, we describe the algorithm in a more detailed way. In Section 3 we evaluate the application of the algorithm to MHS (Microwave Humidity Sounder) instead of AMSU-B data, which is necessary for extending the data set to cover recent years, performing a comparison with different in-situ data sources in Section 3.2 and to ERA5 reanalysis in Section 3.3. Then, in Section 4 we evaluate the new ice cloud filtering developed for the algorithm in Section 4, and finally give some conclusions in Section 5.

2 Retrieval algorithm

2.1 Data sources

The algorithm uses microwave radiometer satellite measurements from humidity sounders such as AMSU-B or MHS on board the NOAA (National Oceanic and Atmospheric Administration) 15 to 19 satellites and Eumetsat EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites) Metop-A, Metop-B and Metop-C satellites. The characteristics of each sensor can be found in Table 1, and the launch dates of each satellite in Table 2. Through this paper, when we refer to AMSU-B TWV, the brightness temperature data used for the retrieval is always from the sensor on NOAA-17, with the version from the Fundamental Climate Data Record (Ferraro and Meng, 2016), which provides an inter-satellite calibrated set of brightness temperatures as described in Ferraro (2016). When we refer to MHS TWV, the brightness temperature data are from NOAA-18, similarly sourced.

Additionally, to distinguish between surface types, the daily ice concentration provided by the ASI-algorithm (Spreen et al., 2008) will be used, with pixels with ice concentrations below 15% as open water, while the ones with more than 80% will be considered ice. The percentages between those will not be used.

2.2 Radiative transfer equation

The algorithm starts from the formulation of the radiative transfer equation in the contracted form by Guissard and Sobieski (1994) which describes the brightness temperature (T_B) measured by a space-borne radiometer as:

$$T_B(\theta) = m_p T_s - (T_0 - T_c)(1 - \epsilon_s)\mathrm{e}^{-2\tau \sec\theta}, \tag{1}$$

where θ is the zenith angle, T_s and T_0 are the surface and air temperatures, respectively, T_c is the cosmic background emission, ϵ_s the surface emissivity, τ_0 the total opacity of the atmosphere in the vertical direction, and m_p a correction to take into account both a non-isothermal atmosphere and the difference between the surface (skin) temperature, T_s, and the temperature of the atmosphere at the ground, T_0, ($m_p = 1$ would be the isothermal case and $T_0 = T_s$). The approach by Melsheimer and Heygster (2008), summarized in the following, assumes the ground to be approximated as a specular reflector, which should be good enough for remote sensing in the frequency range we are dealing with, according to Hewison and English (1999).

2.3 Retrieval for equal emissivity assumption

Note that the entire derivation of the final total water vapour retrieval equation from the radiative transfer equation is described in detail in the initial paper for the Antarctic by Miao et al. (2001) and the subsequent Arctic extension by Melsheimer and Heygster (2008). We summarize it here because the basic mechanism is necessary to understand the changes performed.

We start from microwave radiometer satellite measurements in three different channels i, j, k, such as mentioned in Section 2.1. We assume none of these three channels are saturated, i.e., the sensor is still sensitive to the whole atmospheric column and ground. Additionally, we take the ground emissivity as equal in all three channels (as they see the same footprint, and the emissivity does not vary between the channels), while the water vapour absorption (mass absorption coefficient k (m^2/kg)) is different, with $k_i < k_j < k_k$. Then, the brightness temperature difference of two channels i,j can be expressed as:

$$\Delta T_{ij} \equiv T_{Bi} - T_{Bj} = (T_0 - T_c)(1-\epsilon_s)(\mathrm{e}^{-2\tau_i \sec\theta} - \mathrm{e}^{-2\tau_j \sec\theta}) + b_{ij}, \tag{2}$$

where τ_i is the nadir opacity of the atmosphere at the frequency of channel i, and b_{ij} is a bias related to the term m_p for the channels i and j:

$$b_{ij} = T_s(m_{pi} - m_{pj}), \tag{3}$$

As shown in Melsheimer and Heygster (2008) – Appendix II, the bias can here be approximated as:

$$b_{ij} \approx \int_0^\infty \left[\mathrm{e}^{-2\tau_i(z,\infty)\sec\theta} - \mathrm{e}^{-2\tau_j(z,\infty)\sec\theta}\right] \frac{\mathrm{d}T(z)}{\mathrm{d}z}\mathrm{d}z, \tag{4}$$

where $T(z)$ is the atmospheric temperature profile. Then we take the ratio of what we call compensated brightness temperature differences:

$$\eta_c \equiv \frac{\Delta T_{0ij}}{\Delta T_{0jk}} = \frac{\Delta T_{ij} - b_{ij}}{\Delta T_{jk} - b_{jk}} = \frac{\mathrm{e}^{-2\tau_i \sec\theta} - \mathrm{e}^{-2\tau_j \sec\theta}}{\mathrm{e}^{-2\tau_j \sec\theta} - \mathrm{e}^{-2\tau_k \sec\theta}}. \tag{5}$$

We can express the opacities τ_i as a sum of the atmospheric constituent contributions to them: water vapour (τ_i^w) and oxygen (τ_i^{oxygen}). The latter is negligible for AMSU-B channels near the water vapour line, so if we take water vapour mass absorption coefficients k_i and TWV W:

$$\tau_i = \tau_i^w + \tau_i^{oxygen} \approx k_i W, \tag{6}$$

If we approximate the differences of exponentials by products in (5) and take logarithms, we get:

$$\ln(\eta_c) = B_0 + B_1 W \sec\theta + B_2 (W \sec\theta)^2 \tag{7}$$

The three constants B_0, B_1, and B_2 depend on the mass absorption coefficients for the different channels. The term quadratic in W can be neglected (Selbach, 2003; Miao et al., 2001) which leaves us with an equation linear in W that can then be solved to yield our retrieval equation:

$$W \sec\theta = C_0 + C_1 \ln(\eta_c) \tag{8}$$

where $C_0 = \frac{B_0}{B_1}$ and $C_1 = \frac{1}{B_1}$. They are determined empirically as calibration parameters from simulated brightness temperatures based on radiosonde profiles by a regression analysis, described in more detail below (Section 2.6).

2.4 Extension of the retrieval

Normally, for TWV values above 7 kg/m^2 , saturation occurs at Channel 19 (183.3±3 GHz). To extend the retrieval range above this threshold, another channel is required that is less sensitive to water vapour to take its place in the triplet. This means that a new set of assumptions has to be made about the surface emissivity influence. For AMSU-B, the next channel ”in line” is the one at 89 GHz (Channel 16). Thus, the three channels i, j, k are now the AMSU-B Channels 16, 17 and 20 (89, 150 and 183.31±7 GHz). Because Channel 16 is so far from the other two, we can no longer assume that it has the same surface emissivity as the others. Therefore the retrieval equation needs to be re-derived with the changed premise: $\epsilon_i \neq \epsilon_j = \epsilon_k$. This leaves us with a similar looking retrieval equation:

$$W \sec\theta = C_0 + C_1 \ln(\eta_c') \tag{9}$$

where η_c' is a modified ratio of compensated brightness temperatures:

$$\eta_c' \equiv \frac{r_j}{r_i} \left(\eta_c + C(\tau_j, \tau_k)\right) - C(\tau_j, \tau_k), \tag{10}$$

and $C(\tau_j, \tau_k)$ is defined as

$$C(\tau_j, \tau_k) = \frac{\mathrm{e}^{-2\tau_j \sec\theta}}{\mathrm{e}^{-2\tau_j \sec\theta} - \mathrm{e}^{-2\tau_k \sec\theta}}, \tag{11}$$

Since now there is a dependence on emissivities ϵ_i, or, equivalently, on reflectivities $r_i = 1 - \epsilon_i$, the surface emissivity at 89 GHz needs to be examined. Ideally, the ratio of corresponding reflectivities would be taken for each footprint. However, that is not possible without knowing atmospheric conditions and surface temperature. As an approximation, the emissivity is parametrized, and fixed reflectivity ratios depending on surface types are obtained. This was done for sea ice in Melsheimer and Heygster (2008) and for open water surfaces in Scarlat et al. (2018). The upper limit of this extended retrieval is about 15 kg/m^2. Here, we will use this extended retrieval only over sea ice.

2.5 The ”sub-algorithms”: regime selection

As described through Sections 2.3 and 2.4, three different channel triplets are used for the retrieval, depending on the water vapour amount and the saturation of channels; hence, there are three ”sub-algorithms” or retrieval regimes. Each sub-algorithm reaches its upper retrieval limit when the channel ~~wich~~ which is most sensitive to water-vapour becomes saturated. In the original algorithm formulation by Melsheimer and Heygster (2008), the switch from one sub-algorithm to the next (always starting with the most sensitive one) is done only when the saturation condition,

$$T_{bj} - T_{bk} > 0 \tag{12}$$

is fulfilled. This means that for each satellite footprint, only one of the three sub-algorithms is finally used. As the sub-algorithms have been calibrated independently, the switch from one to the next can cause a jump in the retrieved value. A method avoiding this discontinuity in the retrieval values will be discussed further in the follow-on paper. Additionally, as the switch between regimes is done in the brightness temperature space, this does not correspond to a strict cut-off point in water vapour. In Table 3 we summarize the characteristics of each regime.

2.6 Bias and calibration parameters

Since we ordered the channels by the water vapour sensitivity ($\tau_i < \tau_j < \tau_k$), the difference of exponentials in ΔT_{0ij} and ΔT_{0jk} is negative. Therefore, the first term of the temperature difference increases with increased emissivity from negative values to 0 (reached when $\epsilon = 1$). η_c doesn't depend on ϵ, which cancels on the ratioing. In a plot with ΔT_{jk} as abscissa and ΔT_{ij} as ordinate, for constant W and varying ϵ, this is a straight line with slope η_c (W), running through the bias points (b_{jk}, b_{ij}). Since the biases depend only weakly on W and ϵ, all straight lines for different W run through almost the same point $F = (F_{jk}, F_{ij})$, which is called focal point by Miao et al. (2001) and Melsheimer and Heygster (2008). The focal point F is found by simulating brightness temperatures for a set of different ϵ, with different input atmospheric profiles (including W) from radiosonde data, and surface temperature taken as ground-level atmospheric temperature (which makes the small emissivity dependence of the biases vanish; see Melsheimer and Heygster (2008) - Appendix II). Having determined the focal point, the simulated brightness temperature differences and corresponding TWV values from the radiosonde profiles can be used to get the calibration parameters C_0 and C_1. Thus, together with the two focal point coordinates F_{jk} and F_{ij}, there is a total of four calibration parameters in the retrieval equation which are derived by this regression. The specific values for each viewing angle and regime of AMSU-B sensor are found in Melsheimer and Heygster (2008) – Appendix III. For MHS, all these calibration parameters were recalculated and are shown in Appendix A.

2.7 Filtering ice cloud artefacts

The effect of ice clouds at the AMSU-B frequencies as studied in Sreerekha (2005) is known, and has been used for detecting tropical deep convection (Hong et al., 2005) and for an automated method for finding polar mesocyclones (Melsheimer et al., 2016). The latter method uses the sensitivity of retrieved TWV to convective clouds with high ice content as one of the main signatures of polar lows. In these cases, since cloud ice particles are strong scatterers in the used microwave range, the radiation from below the clouds is scattered strongly and hardly reaches the sensor, so that the AMSU-B retrieval is only sensitive to atmospheric water vapour above such clouds and retrieves erroneously low TWV. A procedure to recognize and screen such cases for the AMSU-B/MHS algorithm has been developed. Cloud ice contents high enough to affect our TWV retrieval are almost entirely caused by strong convective clouds which are typically organised in rather small-scale (tens of kilometres) cells or clusters thereof, or which take the shape of mesoscale structures such a polar lows with extents of at most a few hundred kilometres; even in large scale, synoptic low pressure systems, convective clouds are organised in clusters and lines with the above-mentioned scales of tens to a few hundred kilometres. Therefore, image processing methods that rely on the size and shape of ice cloud artefacts can be used: Our approach for eliminating the affected TWV is to find connected areas – minimum of two pixels – of low TWV (<4 kg/m^2) smaller than 50 pixels which are surrounded by higher or non-retrieved values and remove them with a succession of morphology operations (Gonzalez and Woods, 2007), using the tools for Python described in van der Walt et al. (2014): First a dilation with a 7x7 square structural element, and then a closing with the same size structural element.

3 Evaluation of retrieval~~with MHS data~~

In this section, the performance of the TWV retrieval using MHS data is evaluated in subsection 3.1. Then, the satellite-based retrieval will firstly be compared with in-situ data in subsection 3.2, and secondly to ERA5 reanalysis data in subsection 3.3.

3.1 Comparison between MHS and AMSU-B based retrieval

As shown in Table 1, there are some frequency and polarization differences between AMSU-B and MHS sensors. According to the analysis in John et al. (2012), there are some non-negligible discrepancies between the brightness temperatures of AMSU-B and MHS for the second and fifth channels (17 – 150 GHz – and 20 – 183.31 $\pm$ 7 GHz – for AMSU-B, respectively), due to the differences in frequency, while the differences in polarization seem not to be relevant. That raises the question of whether the TWV algorithm will perform equally when using MHS data as input, and, if that is not the case, which adaptation would be needed to ensure consistency of the retrieval results. One main adjustment we did to the retrieval for MHS is the recalculation of all the calibration parameters as described in Section 2.6 and shown in Appendix A.

First, we evaluate the performance for the retrieval as a whole by comparing the retrieved data of both algorithms in the overlap period of both sensors (2008-2009). For this analysis, we considered all the coincident points in the daily gridded data with a 0.25° grid. Figure 1 shows two density plots for the overlap months of January (top) and July (bottom) of 2008-2009. The results of a least squares regression are shown in the Figure as well. Both data sets show good agreement, with most of the points along the one-to-one line. However, we can observe some outliers with high MHS TWV and low, almost constant, AMSU-B TWV, and vice versa, specially striking during the month of July. ~~They are likely caused by~~ These points are mostly associated with time differences of the satellite overpasses, and amount to only about 0.27~~%~~ % of the data, so they are not significant in the overall picture.

In Table 4, the fit statistics for all months are shown. The correlation ranges from 0.87 in June to 0.94 in September. The lowest slope (0.82~~ kg/m^2~~) is found in December~~, which may be related to the smaller than expected amount of data for this month compared to the coincidences in the other winter months~~. On the other hand, the slope is closest to 1.0 in May (0.91~~ kg/m^2~~). The intercept increases for the summer months (June, July, August) but is relatively small for the other months. The RMSD has a similar behaviour: we find higher values for the central months of the year, with a maximum of 2.25 kg/m^2 in August, coinciding with the increased number of outliers. Minimum is of 0.73 kg/m^2 in March. The bias is generally small (minimum of 0.04 kg/m^2 in March, maximum of 0.49 kg/m^2 in September), and positive except for May and June. In general, all parameters show lowest agreement in the summer months when the atmospheric variability is highest. However, we presume the strongest contribution to the ~~low~~ lower agreement in summer is due to the higher uncertainty and variability in the surface emission due to melt process and occurence of melt ponds.

To check any possible influence from the surface type in the consistency of our retrievals, we have studied the TWV time series during 2008-2009 for MHS and AMSU-B over different surfaces: ice, land and open water. The location chosen for each study point is shown in Figure 2, with the surface classification used in the TWV retrieval for a day in early March 2008 (maximum ice extent) as background. ~~Using the ice concentration provided by the ASI-algorithm (Spreen et al., 2008), pixels with ice concentrations below 15% will be open water, while the ones with more than 80% will be considered ice.~~ We show the monthly and yearly means of this time series for the four different locations in Figure 3. Note the lack of data for summer months over open water and ocean because of the limitations of the algorithm. All four time series show good agreement which confirms the consistency between our retrievals. The bias and RMSD are small for all four surface types (ice: 0.1±0.4 kg/m^2, open water:0.03±0.15 kg/m^2, marginal ice zone: 0.2±0.7 kg/m^2, land: 0.12±0.19 kg/m^2), but slightly higher in two cases with ice surfaces, which agrees with the higher error of our method for higher water vapour values (extended regime).

3.2 Comparison with in-situ data sources

While TWV retrieved from AMSU-B has been validated with different data sources (Rinke et al., 2009; Palm et al., 2010; Buehler et al., 2012), the same cannot be said about the retrieval with MHS data. Therefore, we ~~performe~~ perform a comparison with TWV derived from radiosondes taken during the N-ICE21015 campaign from January to June 2015 onboard research vessel Lance north of Svalbard (Hudson et al., 2017; Cohen et al., 2017). We select the MHS data as the mean of all the values in a 50 km radius around the location of each radiosonde. The resulting time series is shown in Figure 4. The first thing to note is that the MHS series ends at the start of June because, afterward, the ~~water vapour values are too high for the retrieval~~surface in the area is considered mixed according to the criteria described in Section 2.1. However, both data sets show good visual agreement, except that MHS is not able to capture some of the quasi periodic peaks in TWV from N-ICE2015 data set (seen roughly every two weeks in February and March). We have eliminated these nine outliers associated with the quasi periodic peaks in TWV from the following analysis. The scatter plot of all overlapping points of both data sets – with the colour scale representing the month of the campaign – shown in Figure 5, confirms the good agreement.

Additionally, we used Global Positioning System (GPS) and radiosounding (RS) TWV observations during the common 2008-2009 period between the AMSU-B and MHS sensors to evaluate the satellite TWV retrieval. GPS and radiosonde TWV have been measured at the five coastal Arctic stations Alert, Eureka, Ny Ålesund, Resolute and Scorebysund, as shown in Figure 6. These datasets are part of a homogenized ~~year~~ time series. From the GPS data, 1-h average values of local integrated TWV have been computed each 6 hours. The radiosoundings have been performed once or twice per day at the selected sites (00:00 and 12:00 UTC). Further details about processing can be found in Negusini et al. (2016). As for the AMSU-B and MHS TWV values, we selected points fulfilling the data conditions of ±1h from the integrated GPS measurements (00:00, 06:00, 12:00, 18:00UTC) and found in a 50 km radius around the GPS/RS stations. ~~The resulting~~ Additionally, TWV data from ERA5 reanalysis (Copernicus Climate Change Service , C3S) were obtained using the same conditions. The resulting the AMSU-B, MHS, ERA5, GPS and radiosonde time series in Figure 7 present generally consistent patterns and reasonable seasonal evolution, with drier winters and wetter summers. Overall, the datasets have worse agreement during the summer months, mainly due to "spikier" data, i.e. more extreme water vapour values. Due to this pronounced seasonal cycle, we separate the results between summer (April to September) and winter (October to March) in the following analysis. There seems to be a slight wet bias in summer for both satellite-derived TWV with respect to the other datasets.

Scatter plots comparing each satellite dataset dataset (both satellite and reanalysis) with both radiosondes and GPS have been prepared for each season and station. As an example, Figure 8 shows the results for Alert. The correlation coefficients vary between 0.55 to 0.820.88, and the correlations in winter seems to be generally lower. We presume this is just a numerical effect because of the narrower data distribution. The RMSD, in contrast, is higher in summer (as seen in Figure 9). The only difference between both satellite-based retrievals seems to be a smaller number of coincident points between the MHS TWV and the radiosondes TWV (approximately half of the data points).

Figure 9 shows all fit parameters for the five stations, with separated results between summer and winter. There seems to be only little difference between the results from the two satellite-based retrievals, which corroborates our confidence in the MHS-based retrieval. Over the three quality indicating parameters RMSD, bias and correlation coefficient there is even a slight, but consistent advantage for the MHS based retrieval. The bias values are almost all negative, and smaller than the RMSD . In other words, the variability of our data is bigger than the systematic difference between the compared datasets. Additionally, RMSD the RMSD is along usual values for TWV studies at high latitudes (as seen in Palm et al. (2010) Palm et al. (2010) for Ny Ålesund and in Buehler at al. (2012) Buehler et al. (2012) for Kiruna), which reassures us on the quality of satellite-based PW TWV retrievals. The higher RMSD values in the Arctic summer correlate with the higher disagreement in Figure 9 can also be seen at high PW values over 7 kg/m^2 during summer for all methods in Figure 8 (left, top and bottom). Such disagreement can be explained by the change in sub-algorithms as explained in Section 2.5One explanation for the smaller bias and RMSD during winter can be that also the absolute values during winter are small. The reason for a low correlation is likely that the temporal coherence is less pronounced.

When fits like in Figure 8 are performed for all stations for ERA5 versus GPS and radiosondes, the slopes are closer to one in summer (0.99 for GPS and 0.87 for radiosondes in average for all stations) but underestimate data to a higher degree in winter (in average, 0.85 for GPS and 0.76 for radiosondes). The behaviour of the correlation coefficient is similar, averaging 0.9 and 0.92 in summer, and 0.85 and 0.75 in winter, for GPS and radiosondes, respectively. These values are very similar to the averages for the satellite data versus the in-situ data. The RMSD and bias are generally small, but smaller in winter. The average RMSD is 1.89 kg/m^2 in summer, 1.10 kg/m^2 in winter for GPS, and 1.58 kg/m^2 in summer and 1.05 kg/m^2 in winter for radiosondes. The average bias is generally negative for GPS, averaging -0.5 kg/m^2 in summer and -0.02 kg/m^2 in winter, while it is always positive for radiosondes, averaging 0.34 kg/m^2 in summer and 0.17 kg/m^2 in winter.

3.3 Comparison with ERA5 reanalysis

We have compiled all the overlapping daily means of TWV from AMSU-B and ERA5 (Copernicus Climate Change Service , C3S) for the complete months of January (top) and July (bottom) from 2008 to 2009, shown in Figure 10. The results of a least squares regression are shown in the Figure as well. Both data sets show good agreement, with most of the points along or parallel to the one-to-one line. ow AMSU-B TWV values compared to high ERA5 TWV values can be observed in both months, but are more prominent in summer. These are remnants of ice cloud artefacts that were not entirely filtered out.

Table 5 shows the fit statistics for all months. The correlation ranges from 0.71 in June to 0.88 in December. The worst slope (1.6) is found in September. On the other hand, the slope is closest to 1.0 in August (0.97). However, the RMSD has higher values for the central months of the year, with a maximum of 5.9 kg/m^2 in August, coinciding with the increased number of outliers. Minimum is of 1.004 kg/m^2 in March. The bias is generally negative, and shows similar behaviour as the RMSD. In general, all parameters show lowest agreement in the summer months when the atmospheric variability is highest.

4 Evaluation of changes/improvements in the retrieval: Filtering ice cloud artefacts

Figure 11 shows daily averaged TWV maps – with the ice cloud mask filtering (Section 2.7) already applied – for the AMSU-B/MHS algorithm (top and centresecond row), as well as from a different data product based on AMSR-E observations (Wentz and Meissner, 2006) over open ocean (third row) and ERA5 reanalysis daily mean (bottom) in winter (left) and summer (right). The days chosen to represent each season (6 January and 6 July, 2008, respectively) show how a typical retrieval looks like for the respective season. The first thing to notice is the difference in spatial coverage of AMSU-B TWV between winter and summer. In summer, AMSU-B/MHS retrieval is restricted to the drier regions, mostly over sea ice and Greenland (the upper limit of the retrieval is usually about 15 kg/m^2 for sea ice surfaces). In winter, the retrieval is possible over most of the land, open water areas and sea ice. Meanwhile, there is no significant coverage variation shown between seasons for the AMSR-E retrieval: most open water areas are covered. In consequence, the area covered by both methods is smaller in summer, as we can note in the map illustrating the regional coverage – for the same days – of both algorithms in Figure 12 (orange area shows joint coverage). Still, TWV is retrieved in most of the Arctic in both seasons. Another consequence is that in summer the overlap area is small. In this particular example of Figure 12, there is no overlap between both datasets. As for the ERA5 dataset, the agreement with both AMSU-B and AMSR-E is qualitatively good, showing similar patterns, particularly in winter.

To visualize the areas affected by the ice cloud maskartifact, Figure 13 shows different areas of interest before (left) and after (right) maskingfiltering, for different days spaced evenly throughout 2008 (each three months approximately: 6th of January, 2nd of April, 6th of July and 14th of October). These areas have been chosen as representative cases for the season. Most features – small regions of low TWV surrounded by high TWV – are removed, but there are still some small areas of low values of TWV (such as the retrieved regions in the land around 70° W, 62°N Figure 13 (October, bottom right)). Note that these incorrectly retrieved areas are surrounded by grey values which represent water vapour too high to be retrieved with the AMSU-B method (about >7 kg/m^2 over ocean or land surfaces). We confirmed by comparison to the ERA5 atmospheric reanalysis that the remaining high TWV values are within the expected range. Also the high, >14 kg/m^2, TWV values on 6th July 2008 in the Hudson Bay area are in agreement with ERA5.

To show the overall effectiveness of the ice cloud maskfiltering, we have compiled all the overlapping retrieved TWV from AMSU-B and AMSR-E for the complete months of January (top) and July (bottom) from 2006 to 2008, shown in Figure 14. Before filtering for ice cloud artefacts (left), there is a big cluster of data with high AMSR-E values for relatively low AMSU-B values. Those correspond to the values affected by convective clouds with high ice content. Note that the overlap area between AMSR-E and AMSU-B is small (Figure 12) and therefore cloud artefacts make up a large fraction of the overlap data points, particularly in summer. After filtering (Figure 14, right) the AMSU-B retrieval, they are gone. Additionally, the fit performed improves significantly, with the correlation reaching 0.6 in summer and the slope getting much closer to one in winter (0.95, as compared to 0.3). Note also the jump in density of the retrieved TWV values caused by switching between sub-algorithms mentioned above (Section 2.5), most notably near 6 kg/m^2 (Figure 14). Between 7.6% (January) and 11% (March) of the data are masked by the ice cloud filter for winter months, while the percentage is much smaller in the summer months, ranging between 0.18% of the data in August to 3.7% in June. In summer, up to 94% of those values (July) come from the overlap area between AMSU-B and AMSRE, with the average 55.5%. In winter, the values from the overlap area average 11.8%.

5 Conclusions

We provide an updated version of the TWV retrieval algorithm that originally uses as input microwave humidity sounder data from AMSU-B. The updated algorithm, can now also use data from MHS, the successor instrument of AMSU-B, and contains a filter for artefacts caused by convective clouds with high cloud ice content. The improved retrieval performs better when compared to another satellite product and to in situ data.

The coefficients in the retrieval algorithm were adapted for MHS (Appendix A). We have investigated the impact of differences between AMSU-B and MHS on the retrieved TWV and have found the differences to be negligible. This means that a consistent continuous data set for the years 1999 until now can be generated from combining AMSU-B and MHS data. Additionally, the MHS-based TWV data have been compared with radiosonde data from the N-ICE2015 campaign, and the results show good performance for MHS TWV. Both satellite-derived TWV have been compared against GPS and radiosonde data for five Arctic coastal stations during 2008 and 2009, and the results are satisfactory, with averaged correlations for all stations and methods 0.82 in summer and and 0.75 in winter, and RMSD along usual values for TWV studies at high latitudes. The satellite based TWV retrieval also compares well with the ERA5 reanalysis. Some artefacts of not filtered ice clouds remain but overall the correlation with 0.79 and RMSD of 3.01 kg/m^2 show good correspondence.

The filter for ice cloud artefacts performs well as shown by comparison with data from the AMSR-E based algorithm that works over open water. A remaining issue are the jumps of retrieved TWV values between the different retrieval regimes. This can, however, in principle be mitigated by comparing root mean square differences and bias for adjacent TWV regimes, and choosing an optimal regime, i.e., channel combination, for the range of the water vapour column. Where regimes overlap, weighted averages can smooth the transition.

The algorithm described here has an upper TWV limit that restricts retrieval in summer to the central Arctic and Greenland. However, when combining the TWV data retrieved by the algorithm described here with TWV retrieved over open ocean from AMSR-E and AMSR2 – the product by Remote Sensing Systems (RSS) (Wentz and Meissner, 2006) – a nearly complete coverage of the whole Arctic year-round is possible, starting in 2000, which is the overall goal of future work.

Appendix A

The following Tables list the calibration parameters C_0, C_1, F_{jk}, and F_{ij} for the TWV retrieval algorithm for the Arctic and – for the sake of completeness – the Antarctic, for 15 viewing angles that span the range of the viewing angles of MHS, calculated in the same way as the parameters for AMSU-B-based retrieval by Melsheimer and Heygster (2008). The retrieval equation is, from (5) and (8):

$$W \sec\theta = C_0 + C_1 \ln\left[\frac{\Delta T_{ij} - F_{ij}}{\Delta T_{jk} - F_{jk}}\right], \quad \text{(A1)}$$

where $\Delta T_{ij} = T_{b,i} - T_{b,j}$, the MHS channels i, j, k are

- 5 (190.31 GHz), 4 (183.31 $\pm$ 3 GHz), 3 (183.31 $\pm$ 1 GHz) for the low-TWV algorithm,
- 2 (157 GHz), 5 (190.31 GHz), 4 (183.31 $\pm$ 3 GHz) for the mid-TWV algorithm,

and, from equations (10) and (9),

$$W \sec\theta = C_0 + C_1 \ln\left[\frac{r_j}{r_i}\left(\frac{\Delta T_{ij} - F_{ij}}{\Delta T_{jk} - F_{jk}} + 1.1\right) - 1.1\right] \quad \text{(A2)}$$

where i,j,k are 1 (89.9 GHz), 2 (157 Ghz), 5 (190.31 GHz) for the extended algorithm.

The calibration parameters for the Arctic (Tables A1–A3) were derived using radiosonde data from those World Meteorological Organization (WMO) stations in the Arctic that are located on the coast or on islands(29 stations), from the years 1996 to 2002, which amounts to about 27000 radiosonde profiles.

Competing interests. No competing interests are present

Acknowledgements. AMSR TWV data are produced by Remote Sensing Systems and were sponsored by the NASA AMSR-E Science Team and the NASA Earth Science MEaSUREs Program. Data is available at www.remss.com.

We gratefully acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project number 268020496 - TRR 172, within the Transregional Collaborative Research Center "ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms, $(AC)^3$", as well as the support by the project INTAROS (INTegrated Arctic Observation System) funded by the European Union's Horizon 2020 Research and Innovation Programme under GA 727890.

The authors would like to thank the Department of Atmospheric Science, University of Wyoming (http://weather.uwyo.edu/upperair/sounding.html) and the N-ICE2015 campaign (www.npolar.no/en/projects/n-ice2015/) for the radiosonde raw data and the International GNSS Service (ftp://igs.ensg.ign.fr/pub/igs/data/) for the GPS data.

We thank the three reviewers for their helpful comments, which helped to significantly improve the manuscript.

References

Bobylev, L. P., Zabolotskikh, E. V., Mitnik, L. M., and Mitnik, M. L.: Atmospheric Water vapour and Cloud Liquid Water Retrieval over the Arctic Ocean Using Satellite Passive Microwave Sensing, IEEE T. Geosci. Remote, 48, 283–294, https://doi.org/https://doi.org/10.1109/TGRS.2009.2028018, 2010.

Buehler, S. A., Östman, S., Melsheimer, C., Holl, G., Eliasson, S., John, V., Blumenstock, T., Hase, F., Elgered, G., Pscheidt, I., Redl, S., and Steinke, S.: A multi–instrument comparison of integrated water vapour measurements at a high latitude site, Atmos. Chem. Phys., 12, 10 925–10 943, 2012.

Chahine, M.: The hydrological cycle and its influence on climate, Nature, 359, 373–380, 1992.

Cohen, L., Hudson, S. R., Walden, V. P., Graham, R. M., and Granskog, M. A.: Meteorological conditions in a thinner Arctic sea ice regime from winter to summer during the Norwegian Young Sea Ice expedition (N-ICE2015), Journal of Geophysical Research: Atmospheres, 122, 7235–7259, https://doi.org/10.1002/2016JD026034, 2016JD026034, 2017.

Copernicus Climate Change Service (C3S), 2020: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, https://doi.org/10.24381/cds.adbb2d47, 2017.

Dessler, A. E., Zhang, Z., and Yang, P.: Water vapor climate feedback inferred from climate fluctuations, 2003–2008, Geophys.Res. Lett., 35, 2008.

Ferraro, R.: AMSU-B/MHS Brightness Temperature - Climate Algorithm Theoretical Basis Document, NOAA Climate Data Record Program CDRP-ATBD-0801 Rev. 1, http://www.ncdc.noaa.gov/cdr/fundamental, 2016.

Ferraro, R. and Meng, H.: NOAA Climate Data Record (CDR) of Advanced Microwave Sounding Unit (AMSU)-B, Version 1.0. [AMSU-B on board N17 2007-2009], https://doi.org/10.7289/V500004W, 2016.

Francis, J. and Hunter, E.: Changes in the fabric of the Arctic's greenhouse blanket, Environ. Res. Lett., 2, 2007.

Ghatak, D. and Miller, J.: Implications for Arctic amplification of changes in the strength of the water vapor feedback, J. Geophys. Res. Atmos., 118, 7569–7578, 2013.

Gonzalez, R. and Woods, R.: Digital Image Process, Pearson, 2007.

Guissard, A. and Sobieski, P.: A simplified radiative transfer equation fo application in ocean microwave remote sensing, Radio Sci., 29, 881–894, 1994.

Hanesiak, J., Melsness, M., and Raddatz, R.: Observed and modeled growing-season diurnal precipitable water vapor in south-central Canada, J. Appl. Meteorol. Clim., 49, 2301–2314, 2010.

Hewison, T. J. and English, S. J.: Airborne retrievals of snow and ice surface emissivity at millimeter wavelengths, IEEE Trans. Geosci. Remote Sens., 37, 1871–1879, 1999.

Hong, G., Heygster, G., Miao, J., and Kunzi, K.: Detection of tropical deep convective clouds from AMSU-B water vapour channel measurements, J. Geophys. Res., 110, 2005.

Hudson, S. R., Cohen, L., Kayser, M., Maturilli, M., Kim, J.-H., Park, S.-J., Moon, W., and Granskog, M. A.: N-ICE2015 atmospheric profiles from radiosondes [dataset], https://doi.org/10.21334/npolar.2017.216df9b3, 2017.

John, V. O., Holl, G., Buehler, S. A., Candy, B., Saunders, R. W., and Parker, D. E.: Understanding intersatellite biases of microwave humidity sounders using global simultaneous nadir overpasses, J. Geophys. Res., 117, D0230, https://doi.org/10.1029/2011JD016349, 2012.

Jones, P. D., Trenberth, K. E., Ambenje, P. G., Bojariu, R., Easterling, D. R., Tank, A. M. G. K., Parker, D. E., Renwick, J. A., Rahimzadeh, F., Rusticucci, M. M., Soden, B., and Zhai, P.-M.: Observations: surface and atmospheric climate change, Clim. Chang. 2007 Phys. Sci. basis. Contrib. Work. Gr. I to Fourth Assess. Rep. Intergov. Panel Clim. Chang., p. 235–336, 2007.

Kiehl, J. T. and Trenberth, K. E.: Earth's Annual Global Mean Energy Budget, B. Am. Meteorol. Soc., 78, 197–208, 1997.

Melsheimer, C. and Heygster, G.: Improved Retrieval of Total Water Vapor Over Polar Regions from AMSU-B Microwave Radiometer Data, IEEE Trans. Geosci. Rem. Sens., 46, 2307–2322, 2008.

Melsheimer, C., Frost, T., and Heygster, G.: Detectability of Polar Mesocyclones and Polar Lows in Data from Space-Borne Microwave Humidity Sounders, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,, 9, 326–335, 2016.

Miao, J., Kunzi, K., Heygster, G., Lachlan-Cope, T., and Turner, J.: Atmospheric water vapor over Antartica derived from Special Sensor Microwave/Temperature 2 data, J. Geophys. Res., 106, 10 187–10 203, 2001.

Miller, J. R., Chen, Y., Russell, G. L., and Francis, J. A.: Future regime shift in feedbacks during Arctic winter, Geophys. Res. Lett., 34, 7–10, 2007.

Negusini, M., Petkov, B., Sarti, P., and Tomasi, C.: Ground based water vapor retrieval in Antarctica: an assessment, IEEE Trans. Geosci. Rem. Sens., 54, 2935–2948, 2016.

Palm, M., Melsheimer, C., Noël, S., Heise, S., Notholt, J., Burrows, J., and Schrems, O.: Integrated water vapour above Ny–Ålesund, Spitsbergen: a multi–sensor intercomparison, Atmos. Chem. Phys., 10, 1–12, 2010.

Rinke, A., Melsheimer, C., Dethloff, L., and Heygster, G.: Arctic total water vapour: Comparison of regional climate simulations with observations and simulated decadal trends, J.Hydrometeorol., 10, 113–129, 2009.

Ruckstuhl, C., Philipona, R., Morland, J., and Ohmura, A.: Observed relationship between surface specific humidity, integrated water vapor, and longwave downward radiation at different altitudes, J. Geophys. Res., 112, 1–7, 2007.

Scarlat, R., Melsheimer, C., and Heygster, G.: Retrieval of total water vapour in the Arctic using microwave humidity sounders, Atmos. Meas. Tech., 11, 2067–2084, 2018.

Screen, J. A. and Simmonds, I.: The central role of diminishing sea ice in recent Arctic temperature amplification, Nature, 464, 1334–1337, 2010.

Selbach, N.: Determination of total water vapour and surface emissivity of sea ice at 89 GHz, 157 GHz and 183 GHz in the arctic winter, Ph.D. thesis, ser. Berichte aus dem Institut fur Umweltphysik, vol. 21. Berlin, Germany, 2003.

Serreze, M. C., Barry, R., and Walsh, J.: The large-scale freshwater cycle of the Arctic, Bull. Amer. Meteor. Soc., 8, 719–731, 1995.

Serreze, M. C., Barrett, A. P., Slater, A. G., Woodgate, R., Aagaard, K., Lammers, R. B., Steele, M., Moritz, R., Meredith, M., and Lee, C.: The large-scale freshwater cycle of the Arctic, J. Geophys. Res.-Ocean., 111, 2006.

Soden, B., Wetherald, R., Stenchikov, G., and Robock, A.: Global cooling after the eruption of Mount Pinatubo: a test of climate feedback by water vapor, Science, 296, 727–730, 2002.

Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davis, S. M., Sanford, T. J., and Plattner, G. K.: Contributions of Stratospheric Water vapour to Decadal Changes in the Rate of Global Warming, Science, 113, 1219–1223, https://doi.org/https://doi.org/10.1126/science.1182488, 2010.

Spreen, G., Kaleschke, L., and Heygster, G.: Sea Ice Remote Sensing Using AMSR-E 89 GHz Channels, J. Geophys. Res., 113, C02S03, https://doi.org/10.1029/2005JC003384, 2008.

Sreerekha, T.: Impact of clouds on microwave remote sensing, Ph.D. thesis, University of Bremen, 2005.

Stevens, B. and Bony, S.: What are climate models missing?, Science, 340, 1053–1054, 2013.

Trenberth, K. E., Smith, L., Qian, T., Dai, A., and Fasullo, J.: Estimates of the Global Water Budget and Its Annual Cycle Using Observational and Model Data, J. Hydrometeorol., 8, 758–769, 2007.

Trenberth, K. E., Dai, A., Fasullo, J., van der Schrier, G., Jones, D. P., Barichivich, J., Briffa, K. R., and Sheffield, J.: Global warming and changes in drought, Nature Climate Change, 4, 14–22, 2013.

van der Walt, S., Schönberg, J., Nunez-Iglesias, J., Boulogne, F., Warner, J., Yager, N., Gouillart, E., Yu, T., and the scikit-image contributors: scikit-image: Image processing in Python, PeerJ, 2, e453, https://doi.org/10.7717peerj.453, 2014.

Weaver, D., Strong, K., Schneider, M., Rowe, P., Sioris, C., Walker, K., Mariani, Z., Uttal, T., McElroy, C., Vömel, H., Spassiani, A., and Drummond, J.: Intercomparison of atmospheric water vapour measurements at a Canadian High Arctic site, Atmos. Meas. Tech., 10, 2851–2880, 2017.

Wentz, F. and Meissner, T.: AMSR ocean algorithm, Algorithm Theoretical Basis Document (ATBD). Version 2. Report number 121599A-1, Remote Sensing Systems, Santa Rosa, CA, 2006, 2006.

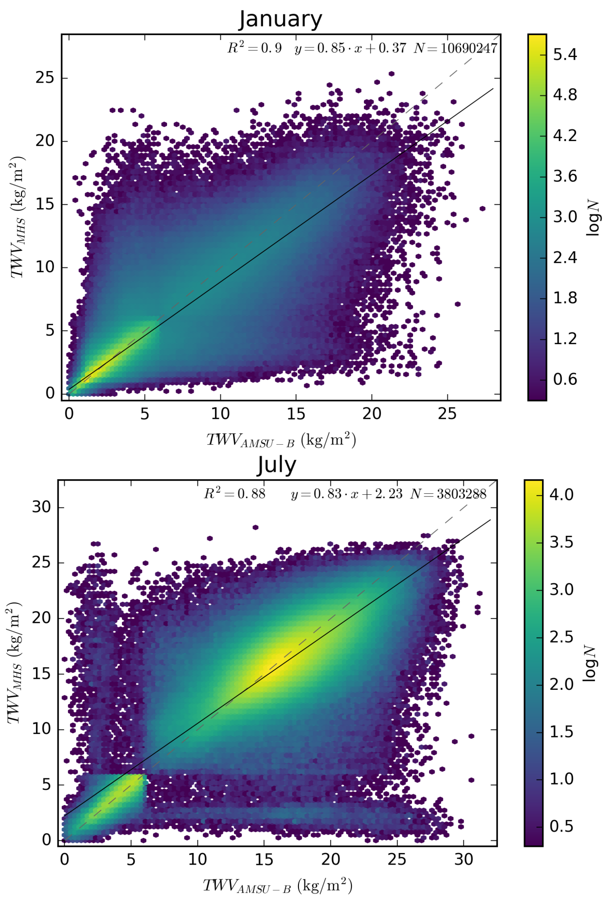


Figure 1. Density plot and fit for MHS TWV versus AMSU-B TWV retrievals for all the coincident points in January (top) and July (bottom), 2008-2010. The dashed line is the one-to-one line, and the black line corresponds to the linear fit of the data.

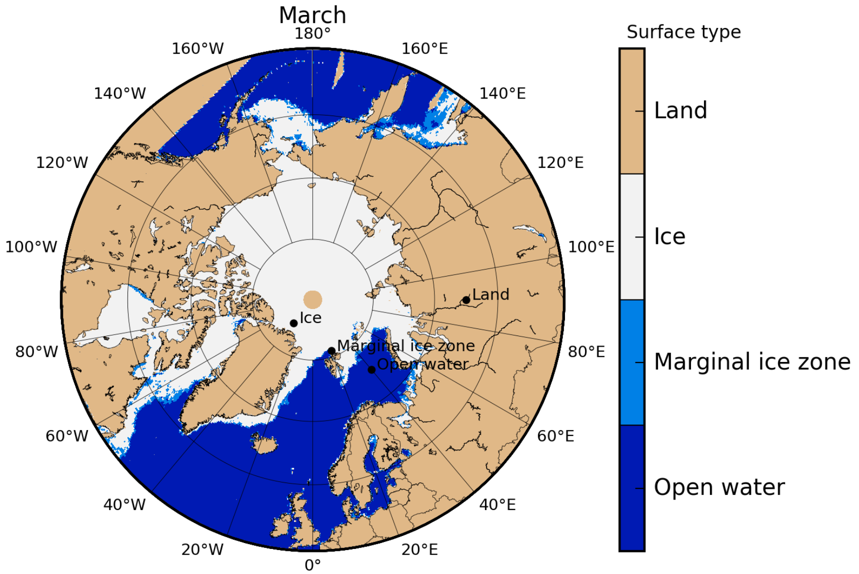


Figure 2. In black, location of the points chosen for the surface characterisation study for TWV. As background, the surface classification used in the TWV algorithm, obtained from ASI algorithm ice concentration (Spreen et al., 2008) for a typical day in March (6.03.2008).

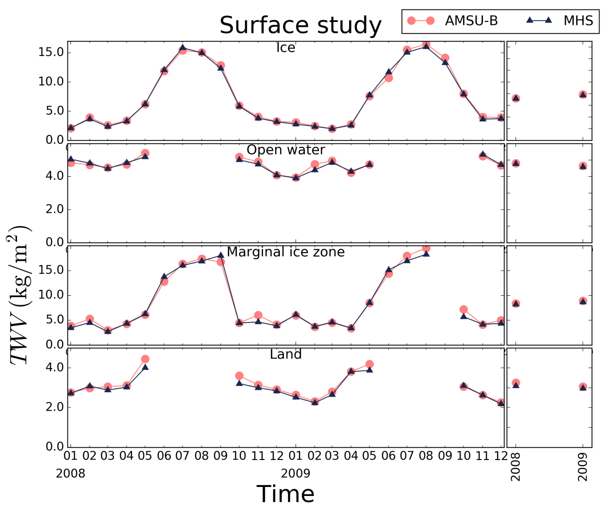


Figure 3. Monthly and yearly means for 2008 and 2009 of the AMSU-B (pink circles) and MHS (blue triangles). TWV retrieval over the locations shown in Figure 2.

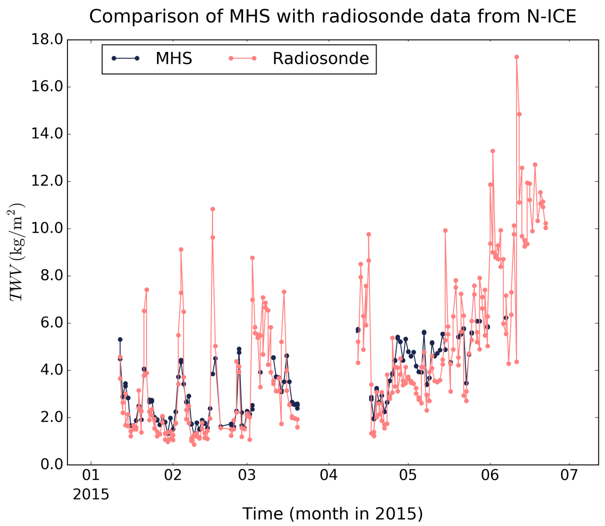


Figure 4. Time series of coincident MHS TWV data (blue symbols) and TWV from radiosondes (red symbols) during the N-ICE campaign.

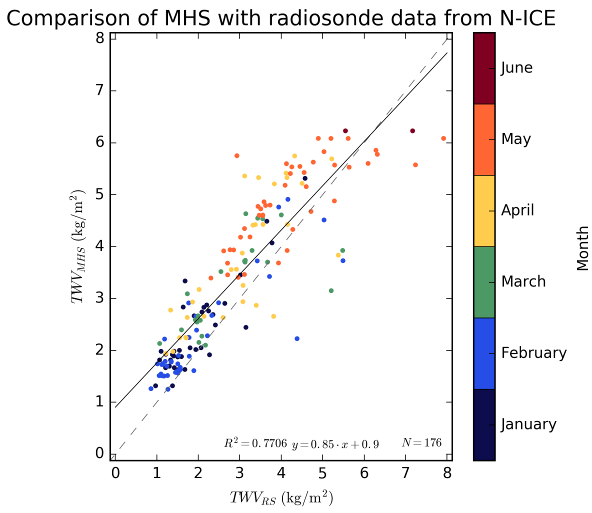


Figure 5. Scatter plot and fit for MHS TWV versus radiosonde TWV retrievals for all coincident points during the N-ICE campaign. The colour scale shows the month where the data point comes from; dashed line: one-to-one lines, solid line: linear regression.

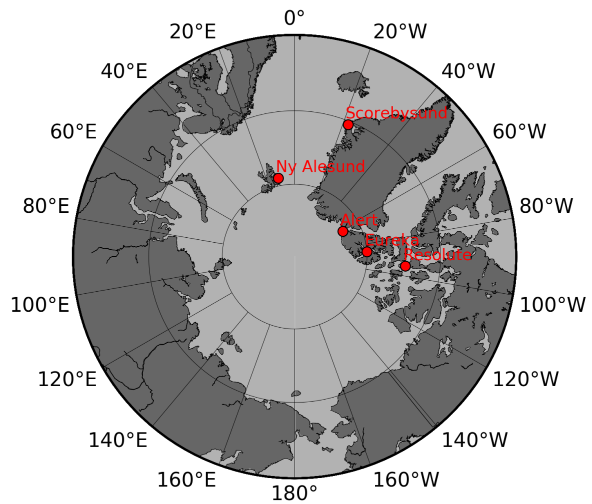


Figure 6. Location of the radiosonde and GPS stations.

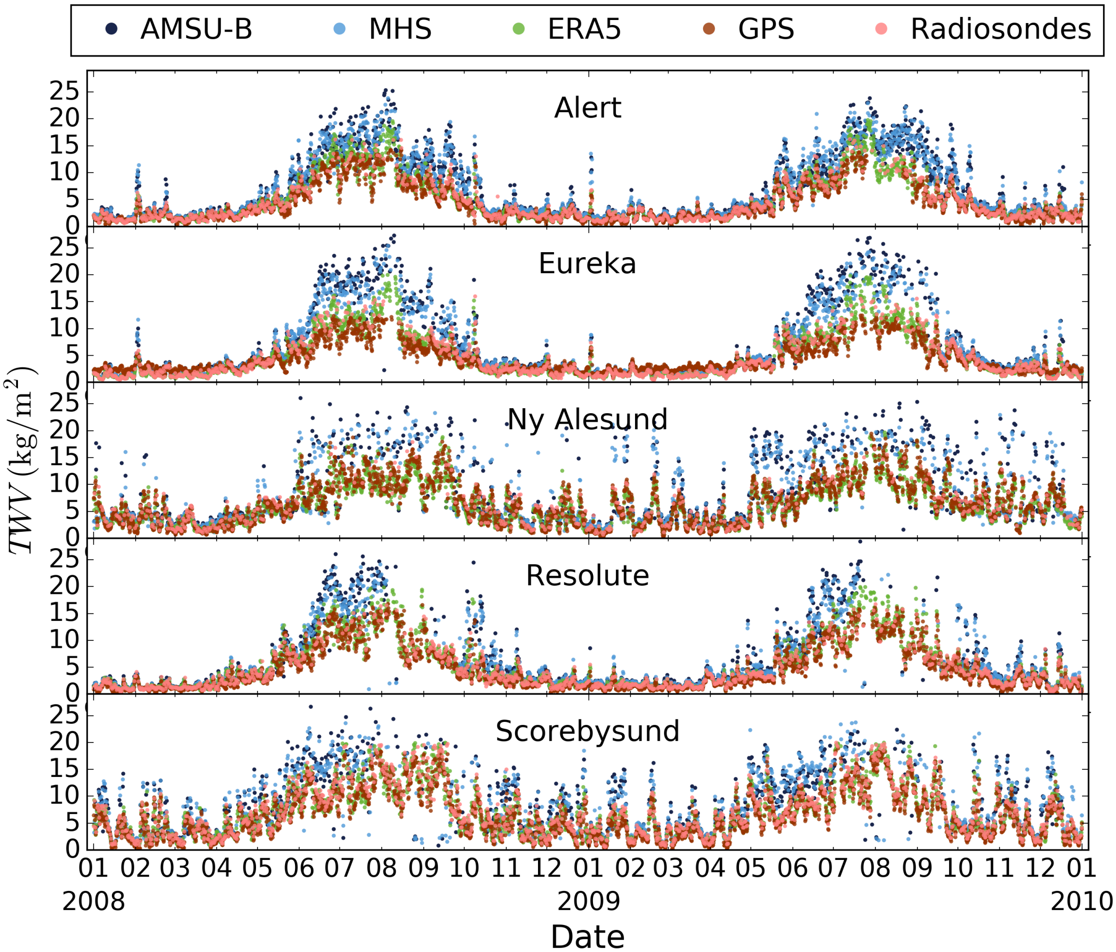


Figure 7. Time series of AMSU-B (dark blue), MHS (light blue), ~~GPS~~ ERA5 (green), GPS (purple) and radiosonde (salmon) TWV retrievals during 2008 and 2009.

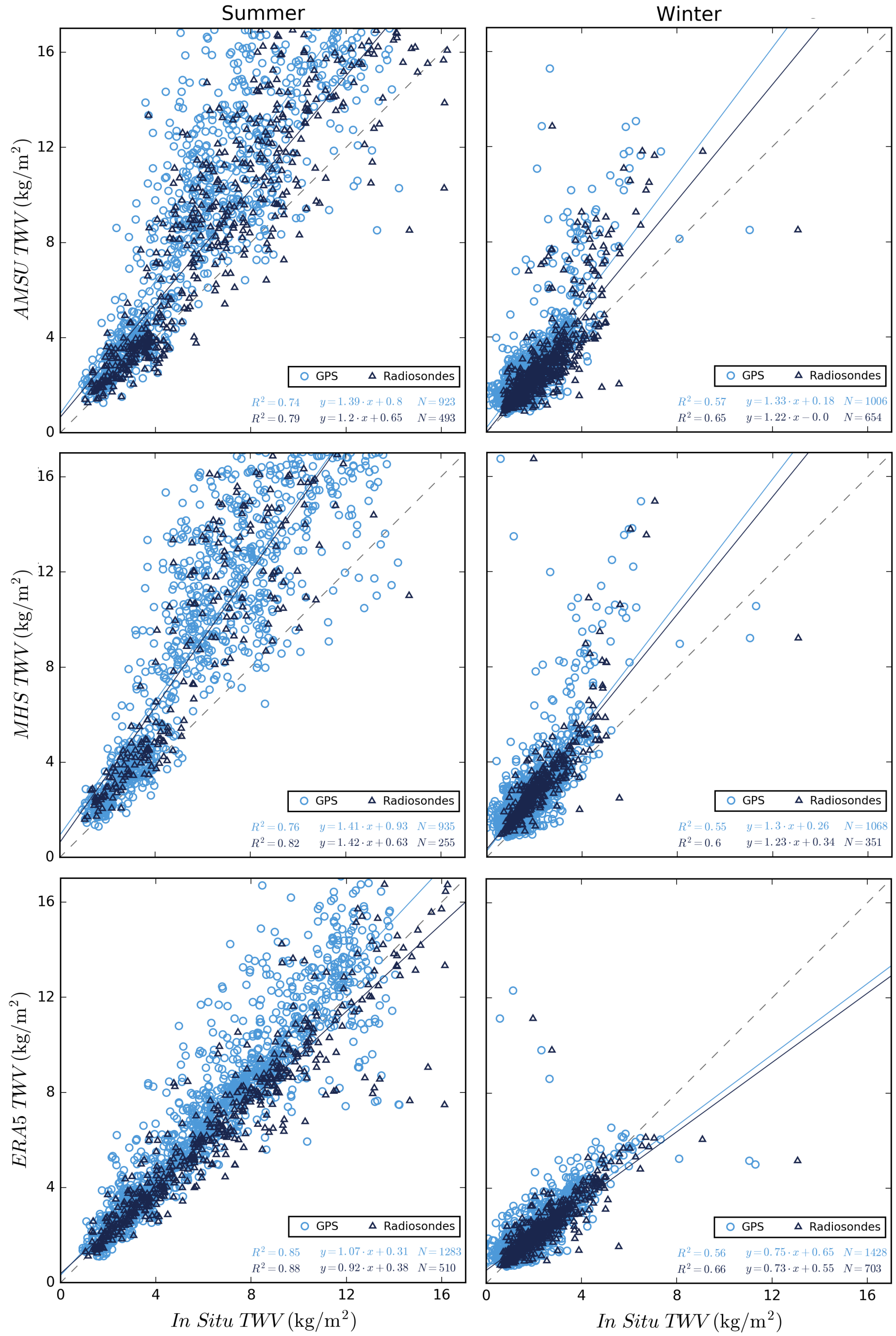


Figure 8. Scatter plots and fits for AMSU-B (top) and MHS (middle) and ERA5 (bottom) TWV retrievals versus GPS (light blue) and radiosondes (dark blue) TWV retrievals for all coincident points during summer (left) and winter (right) of 2008 and 2009 in the Alert station. The solid lines in light and dark blue show the linear regressions for GPS and radiosondes in each case, while the dashed lines are the identity line

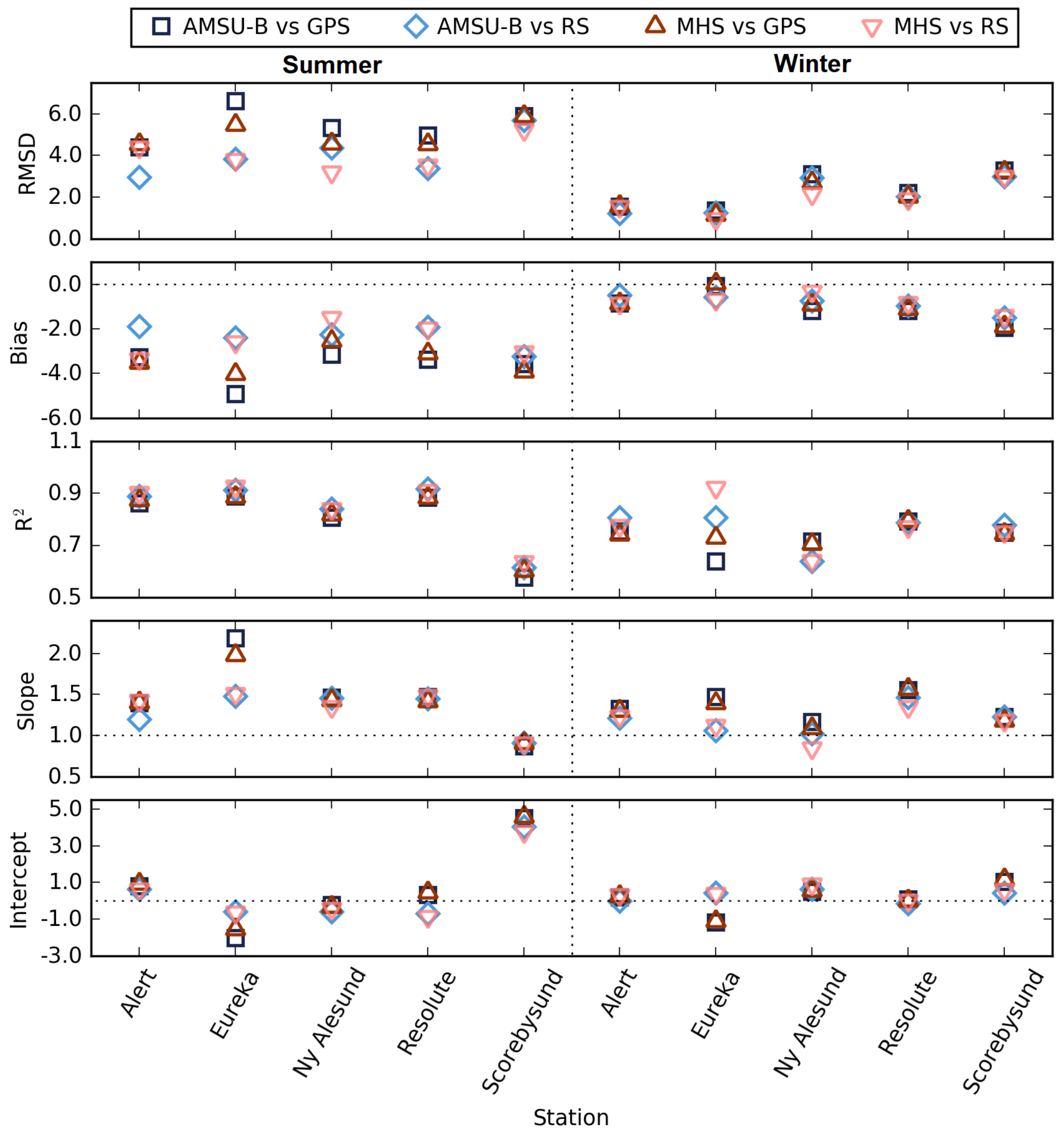


Figure 9. Values of fit parameters for summer (left) and winter (right): RMSD, bias, correlation coefficient R^2, slope and intercept of regression line for MHS and AMSU-B TWV retrievals versus radiosonde and GPS TWV retrievals. RMSD, bias and intercept are in kg/m^2; slope and R^2 are absolute numbers.

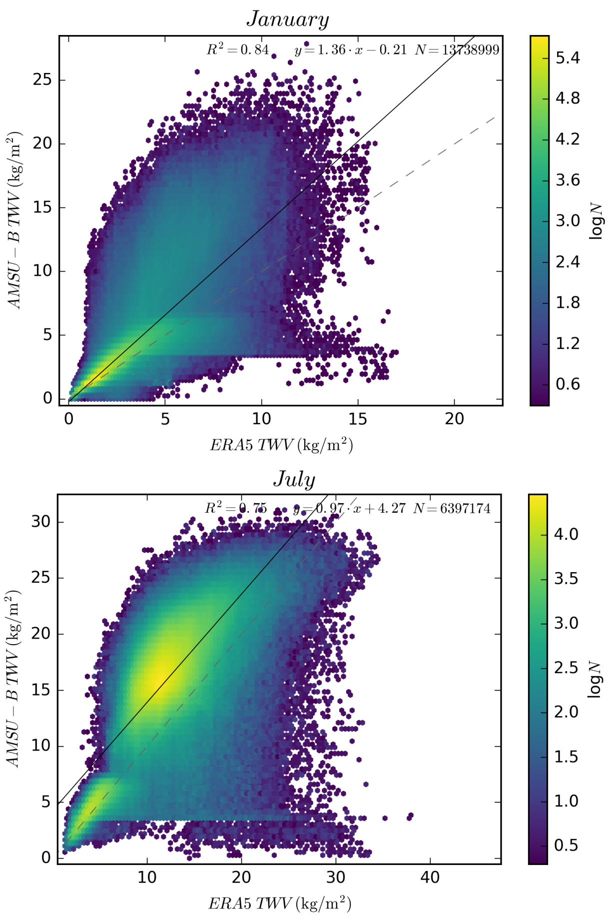


Figure 10. Density plot and fit for AMSU-B TWV versus ERA5 TWV retrievals for all the coincident points in January (top) ~~, MHS (centre)~~ and ~~AMSR-E~~ July (bottom) ~~TWV retrievals for (left) winter (6 January,~~ from 2008 ~~)~~to 2009, ~~and~~ with a fit (~~right~~black solid line) ~~summer~~ for the data clusters over the 1-1 line (~~6 July 2008~~dashed grey).

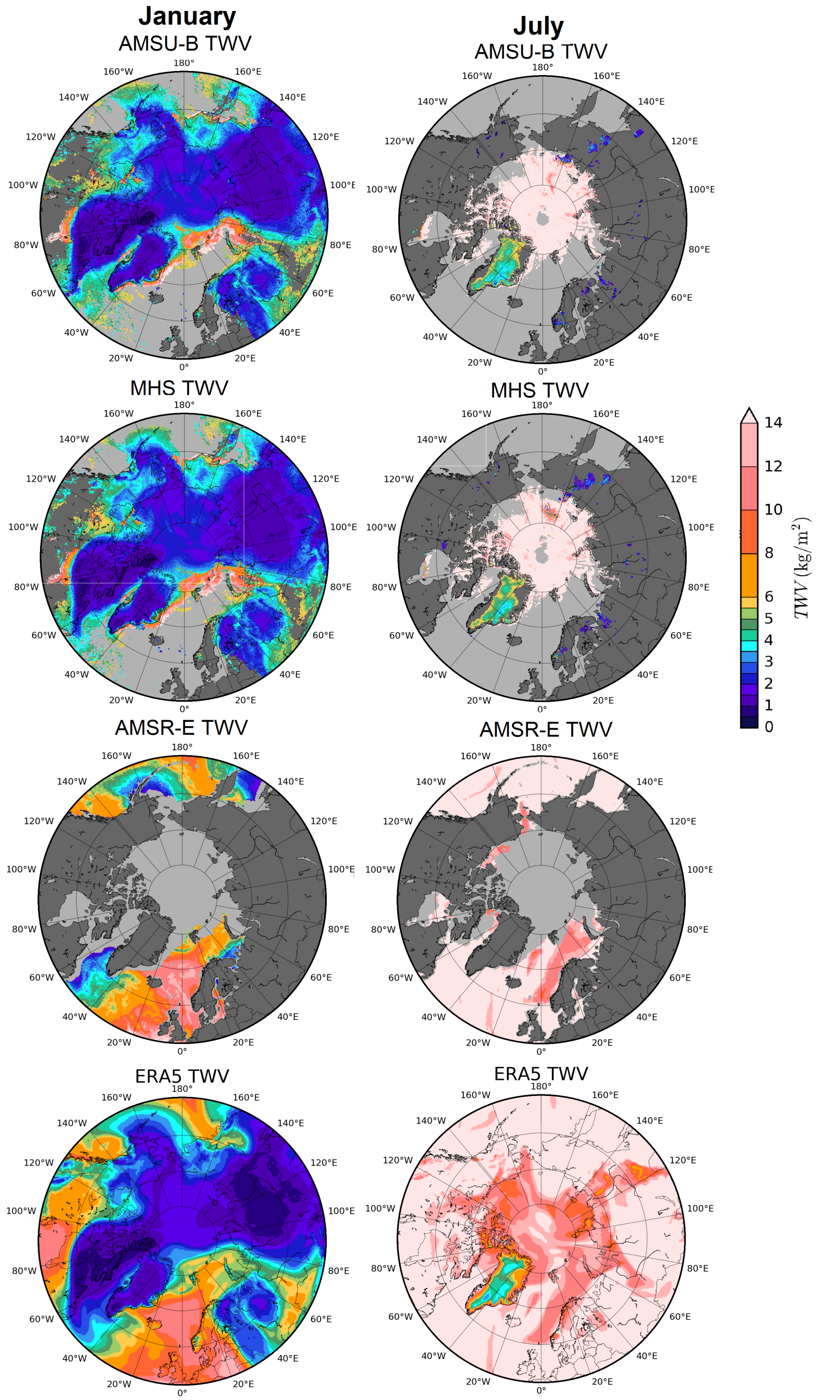


Figure 11. AMSU-B (top), MHS (second row), AMSR-E (third row) and ERA5 (bottom) TWV retrievals for (left) winter (6 January, 2008), and (right) summer (6 July 2008)

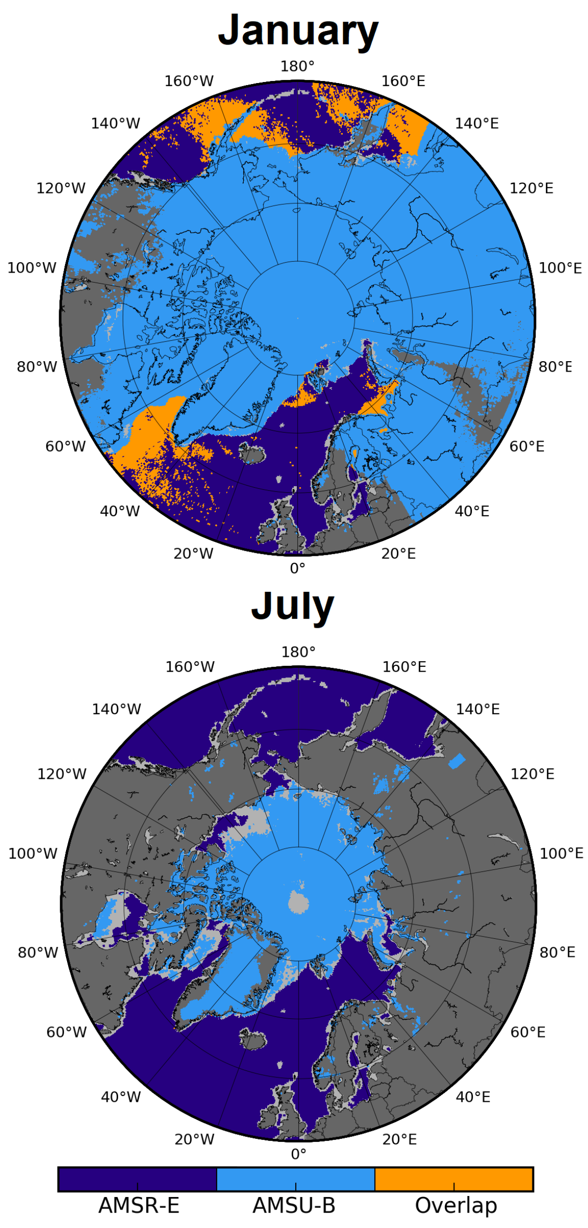


Figure 12. Coverage and overlap area of the merged AMSU-B and AMSR-E retrieval for (top) winter (6 January, 2008), and (bottom) summer (6 July, 2008). Note that there is no overlap between retrievals (orange) for the summer case presented.

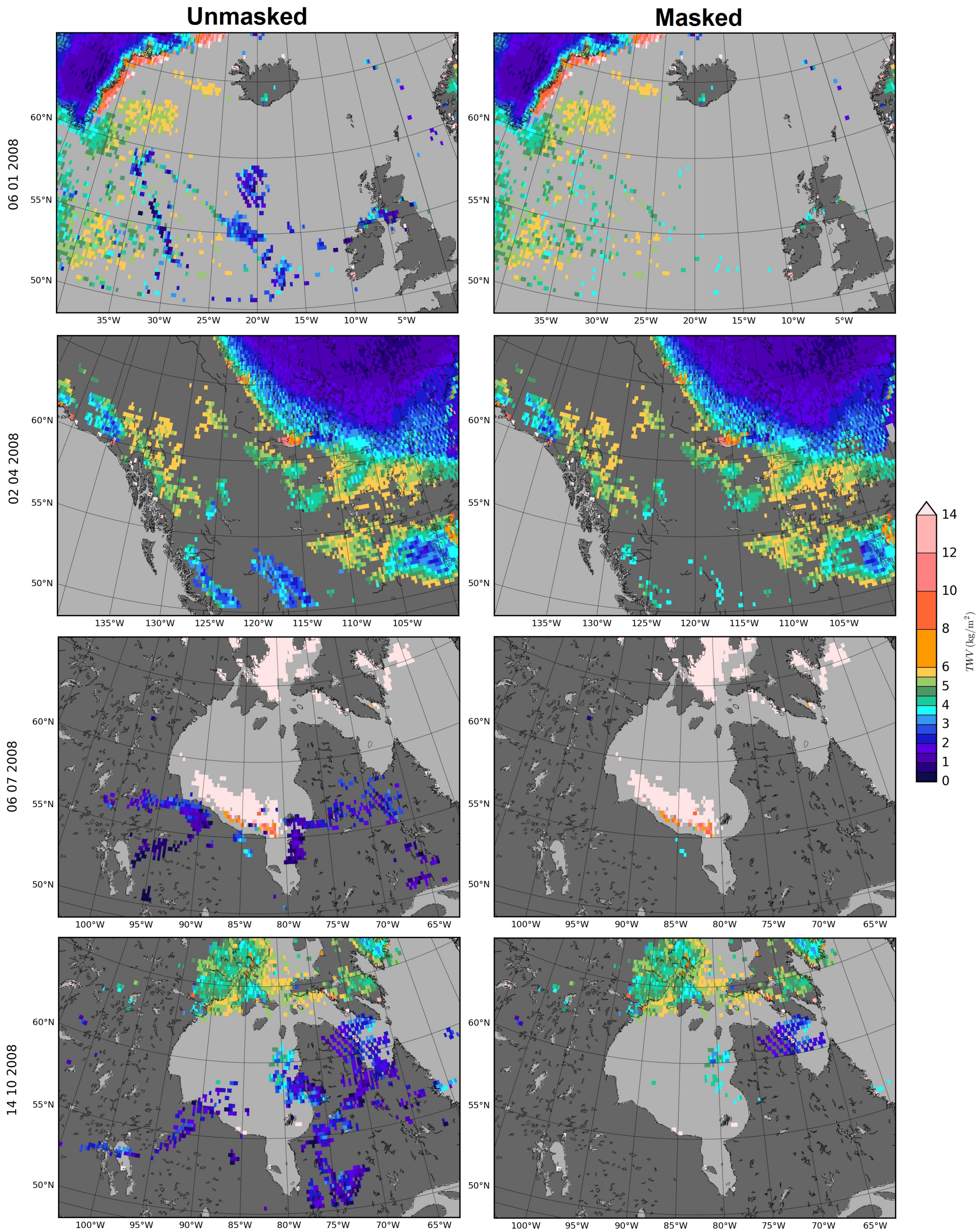


Figure 13. Unmasked (left) and masked (right) AMSU-B TWV retrieval for different showcased areas of four days through 2008: 6 January (top), 2 April (middle up), 6 July (middle down) and 14 October (bottom). Please note the different location in each case.

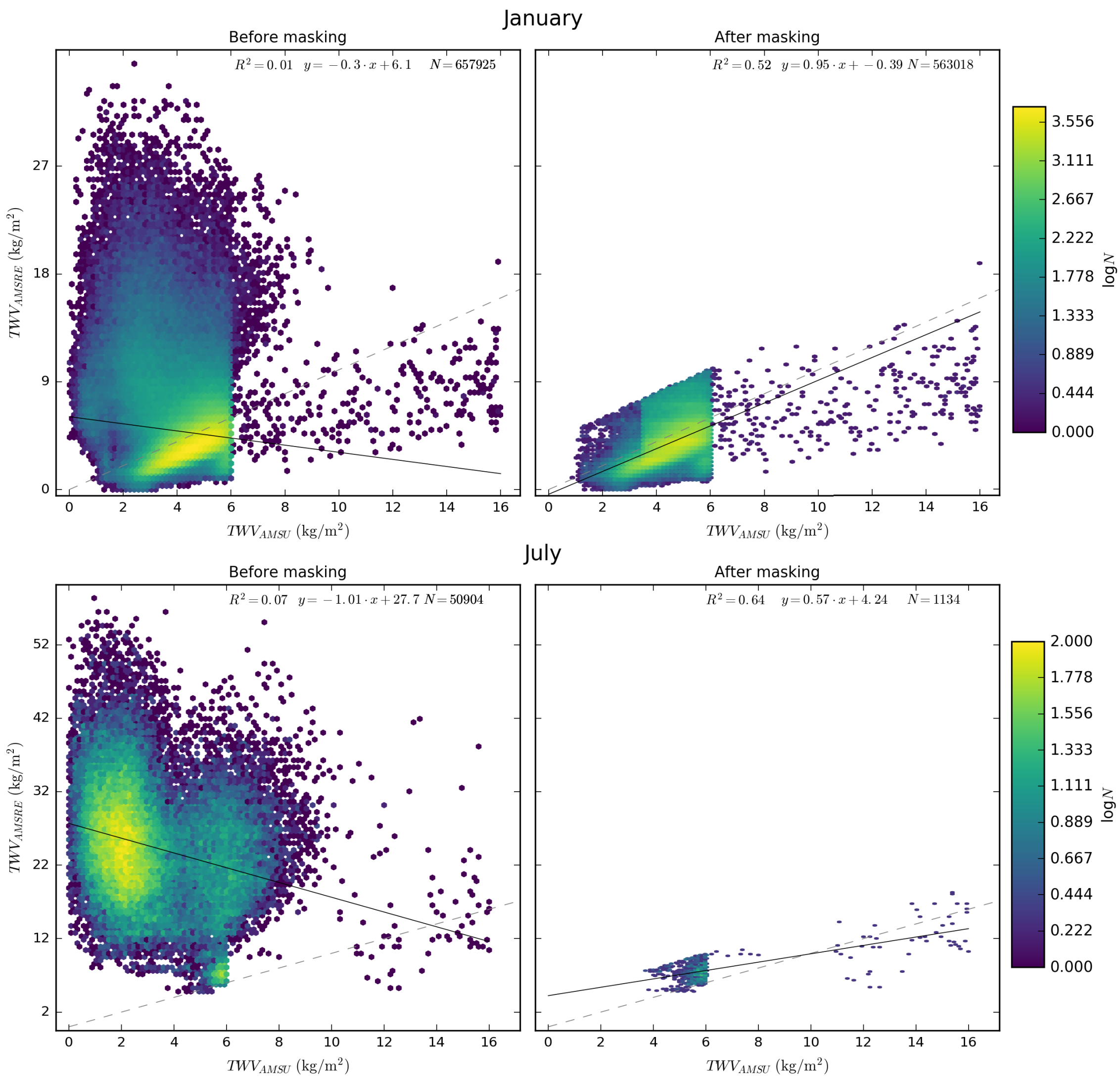


Figure 14. Density plot and fit for AMSR-E TWV versus AMSU-B TWV retrievals for all the coincident points in January (top) and July (bottom) from 2006 to 2008, before (left) and after (right) filtering AMSU-B retrieval for ice cloud artefacts, with a fit (black solid line) for the data clusters over the 1-1 line (dashed grey).

Table 1. Frequency and polarization details for each channel of AMSU-B and MHS sensor

AMSU-B			MHS		
Channel	Frequency (GHz)	Polarisation	Channel	Frequency (GHz)	Polarisation
16	89.9 ± 0.9	Vertical	1	89.9	Vertical
17	150.0 ± 0.9	Vertical	2	157.0	Vertical
18	183.31 ± 1	Vertical	3	183.31 ± 1	Horizontal
19	183.31 ± 3	Vertical	4	183.31 ± 3	Horizontal
20	183.31 ± 7	Vertical	5	190.311	Vertical

Table 2. Humidity Sounders in orbit with platforms, launch year and approximate equator crossing times (ECT) [NOAA]

Platform	Sensor	Launch year	ECT
NOAA15	AMSU-B	1999	07:00
NOAA16	AMSU-B	2000	21:00
NOAA17	AMSU-B	2002	07:00
NOAA18	MHS	2005	20:00
NOAA19	MHS	2009	20:00
MetOp-A	MHS	2006	9:30
MetOp-B	MHS	2012	9:30
MetOp-C	MHS	2018	9:30

Table 3. Characteristics of the different sub-algorithms of the AMSU-B/MHS TWV retrieval. The channel combination is described with AMSU-B frequencies, the MHS retrieval uses the corresponding ones

Sub-algorithm	Chanel combination			Operating surface	Approximate limit TWV (kg/m^2)
Low	183.31 ± 7	183.31 ± 3	183.31 ± 1	Sea ice, ocean, land	1.5
Middle	183.31 ± 7	183.31 ± 3	150	Sea ice, ocean, land	7
Extended	183.31 ± 7	150	89	Sea ice	15

Table 4. Parameters for linear regression for monthly MHS and AMSU-B intercomparison

Month	R^2	Slope	Intercept (kg/m^2)	RMSD (kg/m^2)	Bias (kg/m^2)	Number of points
January	0.90	0.85	0.37	0.97	0.06	10691385
February	0.89	0.84	0.38	0.87	0.05	9858305
March	0.90	0.87	0.31	0.73	0.04	10389349
April	0.90	0.88	0.40	1.02	0.06	8592621
May	0.91	0.91	0.61	1.59	-0.02	6087842
June	0.87	0.84	2.33	2.25	-0.38	4741678
July	0.88	0.83	2.23	2.18	0.38	3803287
August	0.92	0.88	1.43	2.07	0.35	3272951
September	0.94	0.89	0.83	1.77	0.49	3630497
October	0.93	0.86	0.67	1.55	0.19	6000153
November	0.90	0.85	0.53	1.20	0.06	8610697
December	0.88	0.82	0.50	1.24	0.12	7723324

Table 5. Parameters for linear regression for monthly AMSU-B and ERA5 intercomparison

Month	R^2	Slope	Intercept (kg/m^2)	RMSD (kg/m^2)	Bias (kg/m^2)	Number of points
January	0.84	1.36	-0.22	1.43	-0.58	13738999
February	0.85	1.33	-0.2	1.25	-0.52	12626117
March	0.85	1.17	0.06	1	-0.44	14004549
April	0.84	1.16	0.12	1.38	-0.61	11891174
May	0.78	1.25	-0.19	2.72	-1.21	8511599
June	0.71	1.11	1.69	4.32	-2.72	7212136
July	0.75	0.97	4.28	4.89	-3.87	6397174
August	0.73	0.97	5.1	5.91	-4.79	5376590
September	0.81	1.54	0.95	5.83	-4.77	5692249
October	0.74	1.67	-0.24	3.54	-2.11	9281562
November	0.77	1.5	-0.41	2.1	-1.05	11880942
December	0.88	0.82	0.50	1.24	0.12	10822902

Table A1. Calibration parameters, Arctic, low-TWV algorithm

θ	C_0 [kg/m^2]	C_1 [kg/m^2]	$F^L_{4,3}$ [K]	$F^L_{5,4}$[K]
1.667°	0.619	1.05	4.86	4.43
5.000°	0.619	1.05	4.87	4.45
8.333°	0.618	1.05	4.90	4.50
11.667°	0.617	1.05	4.94	4.58
15.000°	0.615	1.05	4.99	4.68
18.333°	0.613	1.05	5.06	4.81
21.667°	0.609	1.05	5.14	4.97
25.000°	0.606	1.04	5.23	5.16
28.333°	0.601	1.04	5.32	5.36
31.667°	0.598	1.02	5.31	5.41
35.000°	0.597	1.00	5.25	5.36
38.333°	0.602	0.96	5.01	4.96
41.667°	0.603	0.92	4.76	4.50
45.000°	0.607	0.87	4.43	3.85
48.333°	0.607	0.80	4.12	3.27

Table A2. Calibration parameters, Arctic, mid-TWV algorithm

θ	C_0 [kg/m^2]	C_1 [kg/m^2]	$F^M_{5,4}$ [K]	$F^M_{2,5}$[K]
1.667°	1.63	2.64	6.56	5.74
5.000°	1.63	2.64	6.55	5.75
8.333°	1.62	2.64	6.54	5.75
11.667°	1.61	2.63	6.52	5.75
15.000°	1.60	2.62	6.50	5.77
18.333°	1.59	2.61	6.46	5.77
21.667°	1.57	2.59	6.43	5.79
25.000°	1.55	2.57	6.38	5.82
28.333°	1.53	2.54	6.34	5.86
31.667°	1.50	2.50	6.25	5.86
35.000°	1.46	2.46	6.18	5.90
38.333°	1.42	2.40	6.09	5.95
41.667°	1.37	2.33	5.99	6.01
45.000°	1.30	2.24	5.83	6.03
48.333°	1.22	2.11	5.65	6.08

Table A3. Calibration parameters, Arctic, extended algorithm

θ	C_0 [kg/m^2]	C_1 [kg/m^2]	$F_{2,5}^E$ [K]	$F_{1,2}^E$[K]
1.667°	14.4	7.45	6.52	0.74
5.000°	14.4	7.47	6.55	0.74
8.333°	14.4	7.50	6.61	0.75
11.667°	14.4	7.56	6.71	0.77
15.000°	14.4	7.63	6.84	0.80
18.333°	14.4	7.73	7.00	0.83
21.667°	14.5	7.83	7.20	0.87
25.000°	14.5	7.97	7.44	0.93
28.333°	14.5	8.11	7.72	1.00
31.667°	14.5	8.26	8.04	1.08
35.000°	14.5	8.43	8.41	1.19
38.333°	14.4	8.60	8.83	1.33
41.667°	14.2	8.76	9.30	1.50
45.000°	13.9	8.90	9.83	1.74
48.333°	13.4	8.99	10.4	2.04