Referee #3

Review of "Assimilation of the AMSU-A radiances using the CESM (v2.1.0) and the DART (v9.11.13)/RTTOV (v12.3)"

This manuscript describes the efforts on assimilation AMSU-A radiances data in the DART system which is coupled with RTTOV123 and CESM. The procedures for quality control, spatial thinning, bias correction and observational errors calculation are developed. The positive impacts are found in the analysis of the primary atmosphere parameters. The paper is well-organized. It appears to be logically set out and the standard of English is acceptable. I recommend this paper for acceptance in the journal with several revisions.

We sincerely appreciate your comments and concerns about this manuscript. Detailed responses to your comments are described below. In this response letter, to distinguish between the referee's comments and authors' responses, your original comments are presented in an Italic font with an underline.

Major comments:

1, Line 150: "the square root of the sum of the observation error variance and the prior background error variance". How about adding a table to list the threshold of each channel?

[Reply]
Thank you for your comment. As mentioned in the manuscript, the threshold value is defined by the square root of the sum of the observation error variance and the prior background error variance. Even though the observation error variance is pre-defined depending on the channel at each satellite platform (shown in Fig. 5), the prior background error variance spatially/temporally varies at each assimilation cycle. That is, the background error variance is derived using the 6-hour forecasts for 20 ensemble members at each analysis cycle. Thus, it is difficult to specify the threshold values employed to filter out the AMSU-A observations whose first-guess departure is large. To make this point clear, we revised this sentence as follows:
[Old, lines 148-151]
"If the difference between the observed AMSU-A brightness temperature and the forward-modeled brightness temperature derived from the model background (6-h forecast) is larger than three times the square root of the sum of the observation error variance and the prior background error variance, the AMSU-A observation is not assimilated (called gross quality control)."
[New]
If the difference between the observed AMSU-A brightness temperature and the forward-modeled brightness temperature derived from the model background (6-h forecast) is larger than three times the square root of the sum of the observation error variance and the prior background error variance, the AMSU-A observation is not assimilated (called outlier test). As the prior background error variance is based on the ensemble spread, the larger the ensemble spread of the 6-h forecast, the more the AMSU-A observations are assimilated."

3, Line 4.2: How was the thinning interval "290 km" determined?

[Reply]
The spatial thinning distance was empirically determined through the extra trial runs. To determine the

optimal thinning distance, we conducted four trial runs in which different thinning distances (i.e., 96 km, 192 km, 288 km, and 384 km) were applied. Among them, the largest analysis benefit was obtained when a thinning distance of 288 km was used. We modified some sentences as follows:
[Old, lines 215-216]
"In this study, the AMSU-A observations are spatially thinned at an interval of about 290 km that was empirically estimated with multiple pre-trial runs."
[Old]
"To choose the optimal spatial thinning distance, we performed four extra assimilation runs in which different spatial thinning distances (i.e., 96 km, 192 km, 288 km and 384 km) were applied. These distances are multiples of the AMSU-A FOV footprint size (~48 km in nadir). The thinning interval of 288 km resulted in the largest analysis impact, so that distance was used to thin the observations in this study."

4, Line 239-242: The AMSU-A data with large latitude (>60°S) is excluded in this paper (Line 204), because their impacts are not ideal. Figure 3(b) shows that, "the residual san biases have different patterns depending on the latitude and for AMSU-A channel 6", and the scan biases are rather large on the latitude band 50S-60S, which is near the >60°S region. It can be expected that the scan biases on the latitude band 60S-90S should be larger. Is it the case? And is it the possible reason for the negative impacts of AMSU-A data >60°S?

[Reply]
Thank you for your comment. As the residual scan biases were estimated using the assimilation outputs (e.g., assimilated observed radiances and background radiances) derived from two-week cycle run, the scan bias patterns seem to be representative of the trial period. And, as you expected, it was a fact that the residual scan bias was large on the high-latitude region (>60°S) in the Southern Hemisphere, as compared with the biases in other latitude regions.
We tried to know why the analysis quality was degraded when the AMSU-A observations in the high-latitude region (>60°S) were assimilated, but it remains unclear. One of the potential reasons is that the bias correction does not work correctly in the high-latitude region, in particular, under extremely cold weather conditions. As mentioned in the manuscript, the scan bias correction considers the characteristics of bias depending on the latitude band (shown in Fig. 3b), but the air-mass bias correction only uses the global-mean bias coefficients. Thus, the global-mean air-mass bias correction coefficients are likely not to work correctly in the high-latitude regions where different model bias pattern occurs locally, in particular, under the extreme weather condition. In fact, the trial experiments in this study were conducted for the summer season (11 August – 30 September 2014) in the Northern Hemisphere. In other words, the high-latitude region (>60°S) including the Antarctic continent was under extremely cold conditions. Thus, we guess that the analysis degradation occurs by applying the global-mean air-mass bias correction that does not consider the bias patterns in the high-latitude region (>60°S). This issue will be handled in the future study.

5, Figure 4: The impact of bias correction on CH10 and CH11 seems not ideal. Especially, the histogram of OMB of CH11 appears to be more "skewed" distribution after bias correction than that before bias correction. And an average deviation of approximately 0.2 K is remained after bias correction. Why the bias correction on CH11 does not perform well? How about adding a table to list the mean value and standard deviations of OMB in each channel before and after bias correction?

[Reply]

First, as the number of O-B samples is relatively small for the AMSU-A channel 11 compared with other channels (shown in Fig. 4 in the manuscript), we update this figure by extending the data period (three weeks from 25 August to 14 September 2014). And, to check the performance of the bias correction process, we add the table in which the mean biases and standard deviations of the first-guess departures are described before and after the bias correction process. For all candidate channels (AMSU-A channels 5–11 for MetOp-B satellite), the mean biases are close to zero, and the standard deviations slightly decrease when the bias correction scheme is applied for the AMSU-A radiances. However, as you mentioned, the distribution pattern of the O-B histogram of AMSU-A channel 11 is slightly skewed after the bias correction compared with other channels showing the Gaussian distribution (Fig. 4). It means that the air-mass bias correction process using the global-mean coefficients is not optimal for the AMSU-A channel 11. It is still unclear why the O-B histogram is skewed for AMSU-A channel 11. Thus, the further study is needed to enhance the performance of the bias correction process.

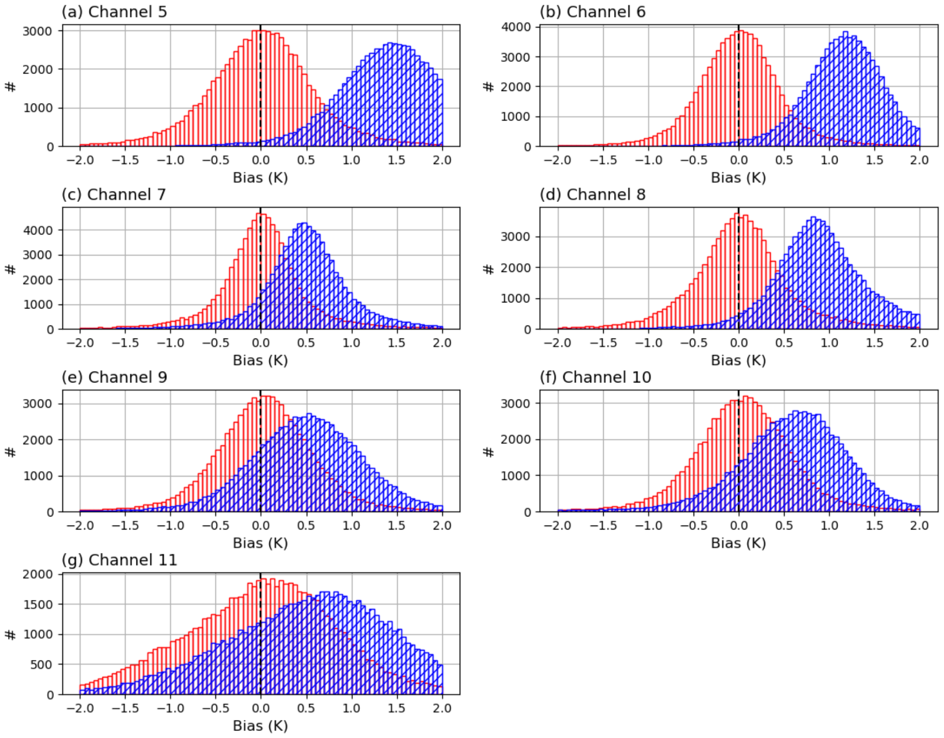

Figure 4. Histogram of the first-guess departures between the observations of the MetOp-B AMSU-A channels 5–11 and the corresponding model background (6-h forecast). Colors indicate the results before the bias correction (hatched blue) and after the bias correction (red), respectively.

Table 2. Mean biases and standard deviations of the first-guess departures (O-B) for MetOp-B AMSU-A channels before and after the bias correction.

O-B	Bias correction	CH5	CH6	CH7	CH8	CH9	CH10	CH11
Bias	X	1.518	1.181	0.514	0.937	0.514	0.590	0.612
	O	0.0005	0.002	0.003	0.014	0.033	0.028	0.010
STDDEV	X	0.677	0.489	0.521	0.572	0.639	0.688	1.052
	O	0.627	0.482	0.494	0.554	0.580	0.642	0.966

6, Figure 6 and Figure 11: The analysis results of temperature, zonal wind and meridional wind is given in detail, while the results of humidity are neither shown in the figures nor mentioned in the text. Although the channels on AMSU-A are mainly sensitive to the temperature, however, as mentioned in Line 382, "a change in one model parameter can change another model parameter in the assimilation process". In my experience, the assimilation of microwave temperature sounders will more or less bring some impacts on the humidity. I wonder the impacts on humidity analysis in this work.

[Reply]
Thank you for your comment. As you mentioned, we computed the error of specific humidity (g/kg or kg/kg) in the analysis against the ERA5 reanalysis to assess the impact of assimilating the AMSU-A observations on the model humidity field. As shown in Fig. S1b, the positive analysis impact (about -5 %) is shown in the troposphere (1000 hPa – 200 hPa) over the Northern Hemisphere where the analysis impact on the primary variables (i.e., 500 hPa geopotential height, temperature, zonal wind, and meridional wind) is significant. In the tropics and Southern Hemisphere, the analysis impact is slightly negative, but not statistically significant, except in the lower stratosphere (near 100 hPa) where the large negative impact (about 18 % and 15 % in the tropics and Southern Hemisphere, respectively) is shown. However, as shown in Fig. S1a, the magnitude of the standard deviation, a denominator to compute the normalized difference of the standard deviation (Fig. S1b), is quite small in the lower stratosphere for the control run. Thus, the change of the standard deviation seems to be negligible in the lower stratosphere. We mention this point in the revised manuscript as follows:
[New]
"In the model humidity field, a positive analysis impact only occurs in the Northern Hemisphere (not shown), but is not as significant as the abovementioned parameters (i.e., 500 hPa geopotential height, temperature, and winds). As a further study, we plan to assimilate the Microwave Humidity Sounder (MHS) providing information on the vertical structure of humidity so that the initial condition of model humidity is improved."

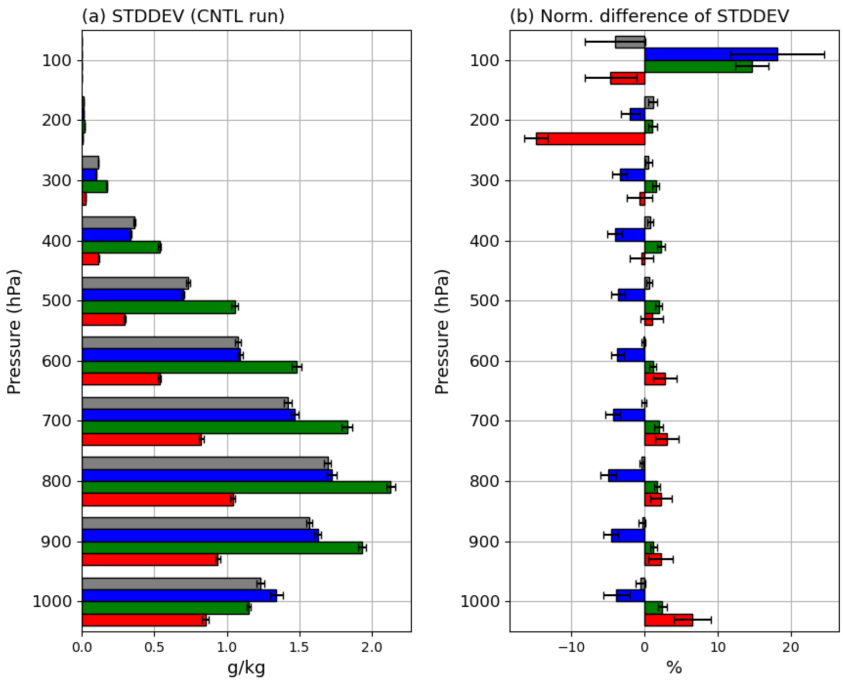

Figure S1. (a) The standard deviation of specific humidity (g/kg) for the control (CNTL) run and (b) normalized difference of the standard deviation of specific humidity between the experiment (AMSU-A) run and the control (CNTL). The 99% confidence intervals are indicated by the horizontal black lines.

7, Figure 11: For both geopotential height, temperature and wind, the impacts of assimilating AMSU-A data in the Northern Hemisphere are better than that in the Southern Hemisphere (Line 444-445, 483-485, 491-493), and these are attributed to the lack of data assimilation in the region >60°S (494-497). However, the observational data from several channels of AMSU-A are also rejected over land (Table 1), which are mostly distributed in the Northern Hemisphere. In another word, the total amount of AMSU-A data assimilated in the Southern Hemisphere should be still more than those in the Northern Hemisphere. In general, the assimilation of satellite data brings more benefits to the analysis and forecasting of Southern Hemisphere, because of the larger ratio of ocean area and the lack of conventional observations. How to understand the difference between the results in this paper and our expectation?

[Reply]
I agree with your opinion. It is well known that the satellite assimilation impact is generally larger in the Southern Hemisphere than its impact in the Northern Hemisphere where the conventional observations are plentiful. In addition, many observations of three tropospheric channels (channels 5–7) are only assimilated over the ocean, which covers more than half of the surface in the Southern Hemisphere. However, in this study, the analysis error reduction is similar in both hemispheres if the high latitude regions (>60°N and >60°S) are not considered in the error computation (see the parentheses in Table 2). It is still questionable why the satellite impact on the analysis is more significant in the Northern Hemisphere. One of the potential reasons is the absence of anchor observations (e.g., radiosonde and GPS-RO) in the Southern Hemisphere, which prevent a drift of the numerical model to its own climatology and biases (Cucurull et al., 2014). In the Southern Hemisphere, where radiosonde observations are rare, the GPS-RO observations are mainly used as the anchor observations. However, in this study, the GPS-RO observations were not assimilated in the trial experiment runs. We also attempted to include the GPS-RO observations in the baseline observations that were assimilated in the control run, because the DART package has the modules for these data. However, it was found that the GPS-RO observations did not provide the analysis benefits. We concluded that additional studies (e.g., error estimation and quality control) are needed to assimilate the GPS-RO observation, thus extracted these data from the baseline observations. In the future study, we will handle this issue.

[Old, line 431]
Table 2. Error statistics of 500 hPa geopotential height (m) for the control (CNTL run) and experiment (AMSU-A run) run. Better values are bolded.

Trial Name	Bias				STDDEV			
	Global	NH	TR	SH	Global	NH	TR	SH
CNTL	-18.70	**-13.90**	-19.05	-27.45	48.82	48.02	13.55	62.54
AMSU-A	**-18.59**	-17.39	**-17.73**	**-25.51**	**42.42**	**31.55**	**12.41**	**58.29**

[New]
Table 2. Error statistics of 500 hPa geopotential height (m) for the control (CNTL run) and experiment (AMSU-A run) run. Better values are bolded. In parentheses, error statistics are shown over the mid-latitude region (30°S-60°S and 30°N-60°N) in the Northern and Southern Hemisphere.

Trial	Bias	STDDEV

Name	Global	NH	TR	SH	Global	NH	TR	SH
CNTL	-18.70	**-13.90** (-18.43)	-19.05	-27.45 (-19.84)	48.82	48.02 (26.71)	13.55	62.54 (38.55)
AMSU-A	**-18.59**	-17.39 **(-16.95)**	**-17.73**	**-25.51** **(-19.54)**	**42.42**	**31.55** **(20.24)**	**12.41**	**58.29** **(33.49)**

Cucurull, L., Anthes, R.A., Tsao, L.-L.: Radio Occultation Observations as Anchor Observations in Numerical Weather Prediction Models and Associated Reduction of Bias Corrections in Microwave and Infrared Satellite Observations, J. Atmos. Ocean. Tech., 31, 20–32, doi: 10.1175/JTECH-D-13-00059.1, 2014.

Minor comments:

1, Figure 2 and Figure 6: the figures should be located behind the paragraph which first mentions it, i.e., behind Line 200 and 372, respectively (unless this manuscript is edited by LaTeX).

[Reply]
Thank you for your comment. We will relocate these figures in the revised manuscript.

2, Line 170-205: The authors describe the quality control procedures as three parts: gross quality control, channel selection, and spatial thinning (Line 15 and 506). However, generally speaking, the contents in Line 170-205 cannot be summarized by "channel selection", because these criterions are applied to the pixels instead of the whole channel. Besides, the spatial thinning is not belonged to quality control, because some pixels are rejected in this procedure not because their quality is not good. Thus, these paragraph should be reorganized. A simple consideration is to replace the title of section 4.1 by "quality control", and revise the corresponding statements in the abstract and conclusion.

[Reply]
Thank you for your comment. We will reflect this point in the revised manuscript.

3, Line 197: "Figures 2a and b" à "Figure 2a and b".

[Reply]
We revise this sentence as follows:
[Old, line 197-198]
"As an example, Figures 2a and b present the spatial distribution of the CLW and the SII retrieved from AMSU-A on board NOAA-19 on 12 August 2014."
[New]
"As an example, Figure 2a and b present the spatial distribution of the CLW and the SII retrieved from AMSU-A on board NOAA-19 on 12 August 2014."

4, Throughout the whole manuscript, sometimes "Figure" is worded, but sometimes "Fig." is worded. Please check the manuscript and unify it.

[Reply]
Following the GMD policy, if the figure is mentioned at the beginning of the sentence, "Figure" is

worded. However, if the figure is referred to in the middle of the sentence, “Fig.” is worded. For this reason, two words (i.e., “Figure” and “Fig.”) are used in the manuscript depending on the location where the figure is mentioned in the sentence.

5, Line 307: “In the pre-trial run, the instrument noise errors were initially used as the observation errors within DART.” How long is the pre-trial run which is used for the statistics of observation errors?

[Reply]
To estimate the observation error variance of AMSU-A channels onboard four satellite platforms, the pre-trial run was conducted for the period from 11 August to 30 September 2014. Except the spin-up period (11 August – 24 August 2014), the background innovations (O-B) and analysis innovations (O-A) were obtained from the output for the pre-trial run. And then, the observations error variances were computed using the estimation method suggested by Desroziers et al. (2005). To make this point clear, we revise this sentence as follow:
[Old, lines 305-307]
“To compute the observation error variances of AMSU-A channels on board four satellite platforms (i.e., Aqua, NOAA-19, MetOp-A, and MetOp-B), the background and analysis innovations were derived from the pre-trial run.”
[New]
“To compute the observation error variances of AMSU-A channels on board four satellite platforms (i.e., Aqua, NOAA-19, MetOp-A, and MetOp-B), the background and analysis innovations were derived from the pre-trial run that was conducted from 25 August to 30 September 2014.”

6, Line 327: “CTRL” à “CNTL”.

[Reply]
Thank you for your comment. We revise this sentence as follow:
[Old, liens 325-327]
“the AMSU-A observations from four LEO satellite platforms (i.e., Aqua, NOAA-19, MetOp-A, and MetOp-B) were assimilated as well as the conventional data that were assimilated in the CTRL run.”
[New]
“the AMSU-A observations from four LEO satellite platforms (i.e., Aqua, NOAA-19, MetOp-A, and MetOp-B) were assimilated as well as the conventional data that were assimilated in the CNTL run.”

7, Line 379: “Figs. 6b and c” à “Fig. 6b and c”.

[Reply]
Thank you for your comment. We revise this sentence as follow:
[Old, lines 378-379]
“In addition to the radiosonde temperature, the first-guess departure errors decrease for the two wind components (i.e., zonal and meridional winds) (Figs. 6b and c).”
[New]
“In addition to the radiosonde temperature, the first-guess departure errors decrease for the two wind components (i.e., zonal and meridional winds) (Fig. 6b and c).”

8, If it is possible, all the figures are better to be parachromatism-friendly. The bars in Figure 4 and 11 are suggested to be shaded by different patterns, just like Figure 8. The symbol should be distinguished *not only by colors but also by shapes (squares, circles, triangles…) in Figure 5, 7, and 9.*

[Reply]

We update the figures as follows:

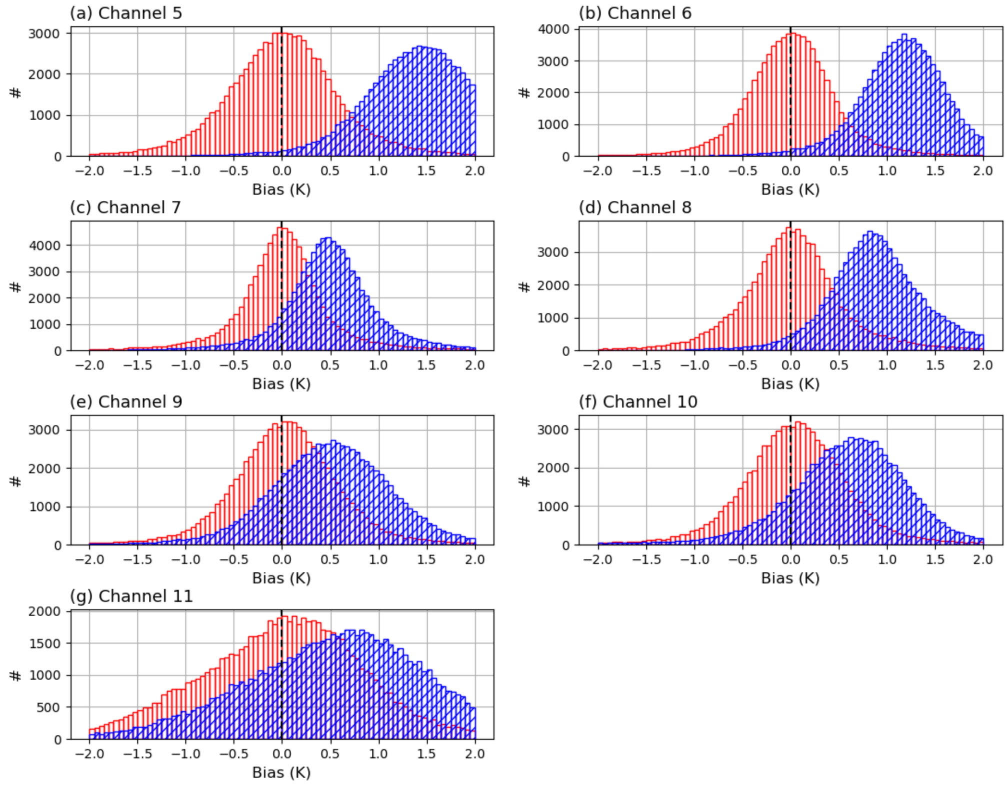

Figure 4. Histogram of the first-guess departures between the observations of the MetOp-B AMSU-A channels 5–11 and the corresponding model background (6-h forecast). Colors indicate the results before the bias correction (hatched blue) and after the bias correction (red), respectively.

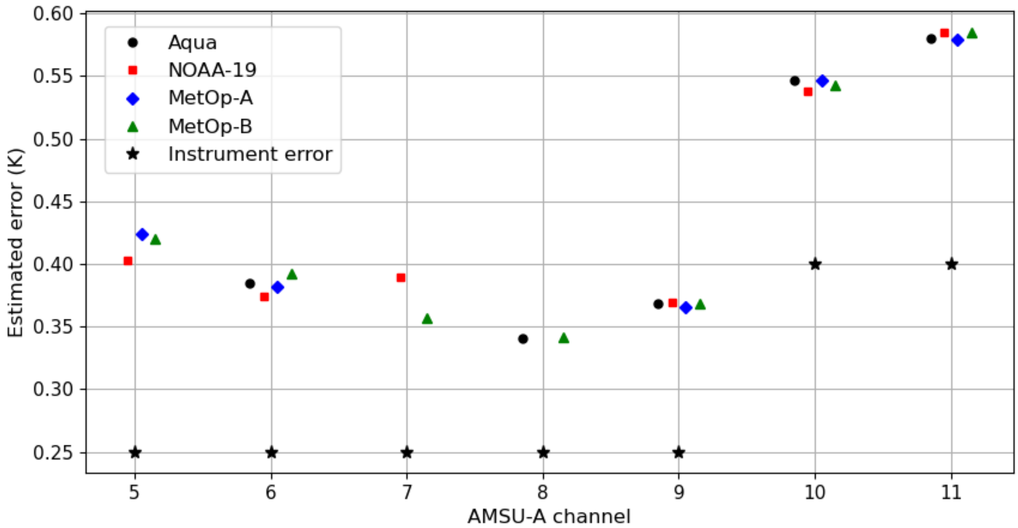

Figure 5. Estimated observation errors (K) for AMSU-A channels on board Aqua (black; circle), NOAA-19 (red; square), MetOp-A (blue; diamond), and MetOp-B (green; triangle) satellite platforms. Black asterisks indicate the instrument noise errors for AMSU-A channels.

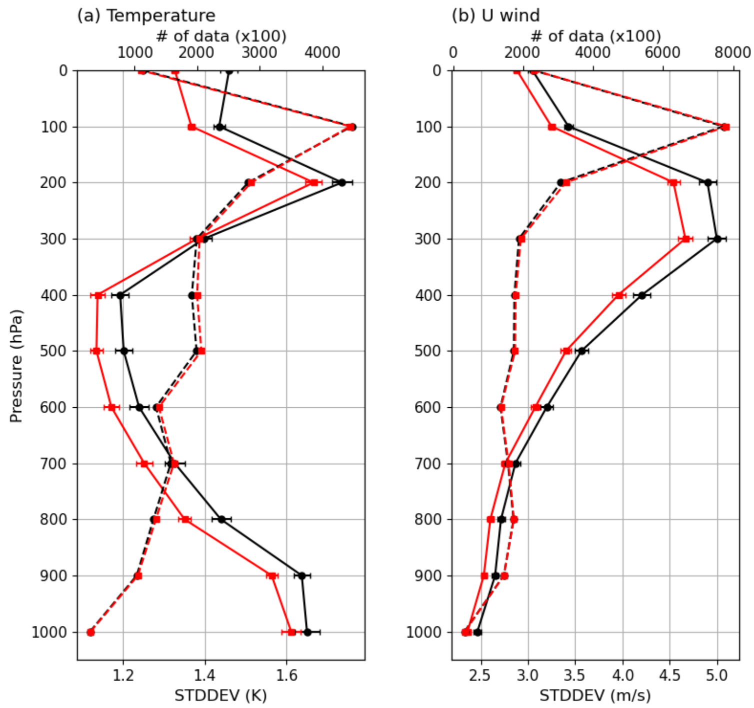

Figure 6. The standard deviation (STDDEV) of the first-guess departures for the radiosonde (a) temperature, (b) zonal wind, and (c) meridional wind for the control (CNTL run; circle symbol and black line) and experiment (AMSU-A run; square symbol and red line) runs. Solid and dashed lines indicate the STDDEV and the number (top axis) of radiosonde measurements assimilated, respectively. The 99% confidence intervals are indicated by the horizontal lines.

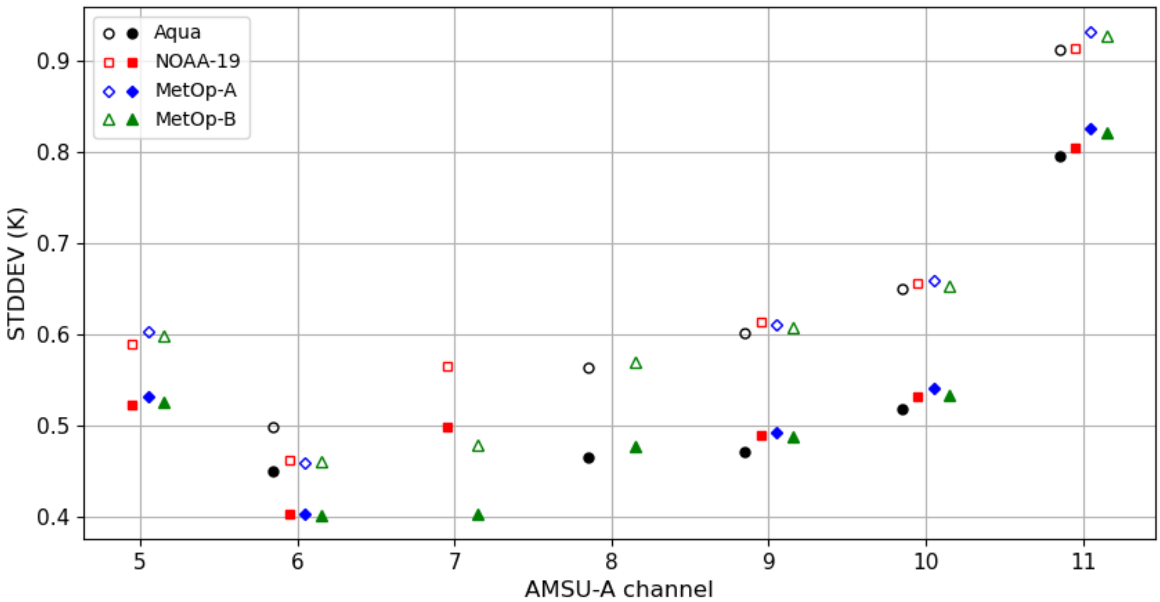

Figure 7. The standard deviations (STDDEVs) of the first-guess departure (unfilled symbols) and analysis departure (filled symbols) for AMSU-A channels on board Aqua (black; circle), NOAA-19 (red; square), MetOp-A (blue; diamond), and MetOp-B (green; triangle) satellites.

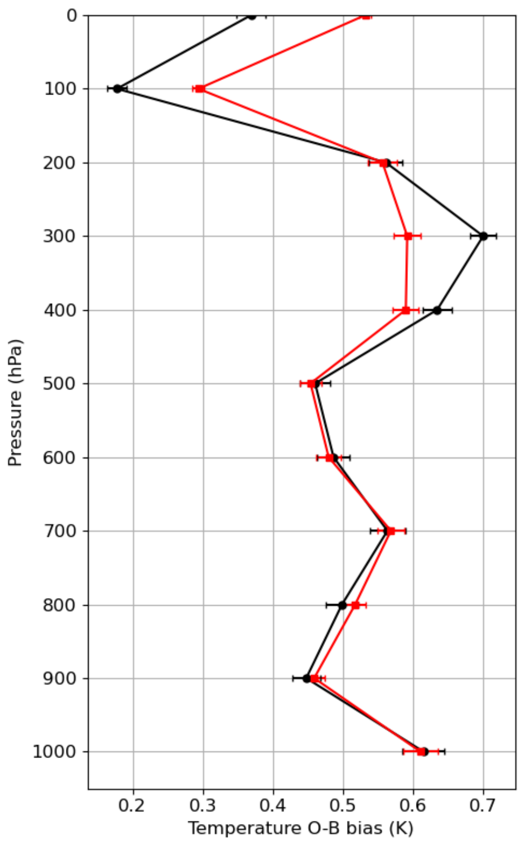

Figure 9. Mean bias of the first-guess departure for the radiosonde temperature measurements for the control (CNTL run; circle symbol and black line) and experiment (AMSU-A run; square symbol and red line) runs. Horizontal lines indicate 99% confidence intervals.

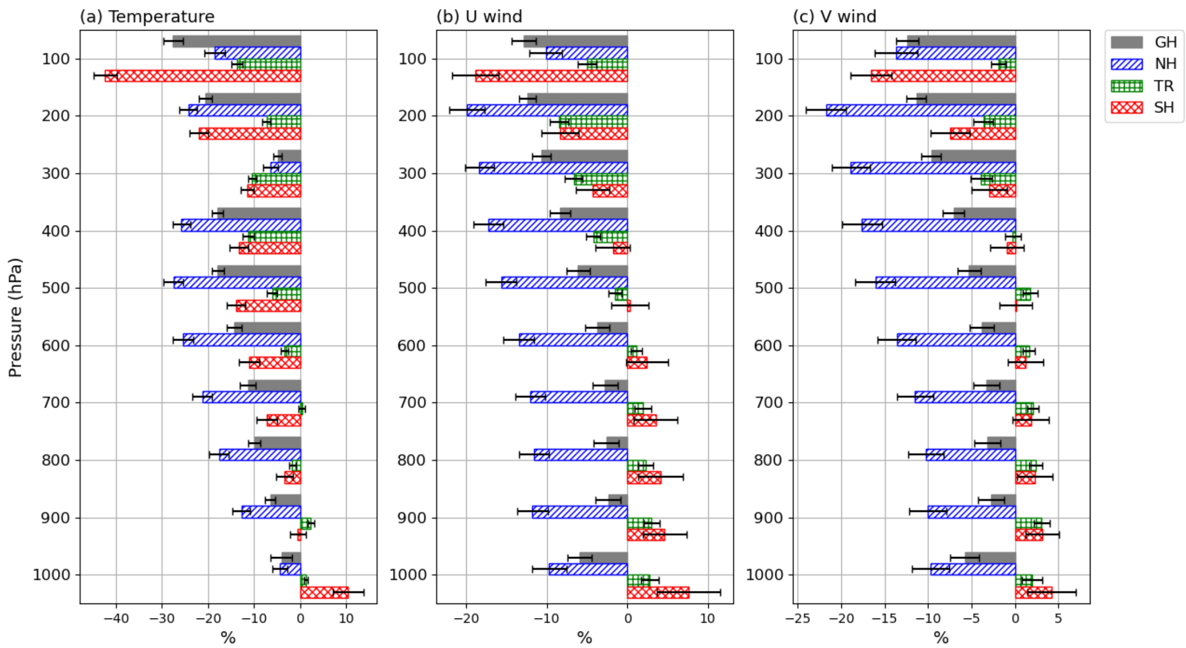

Figure 11. Normalized difference of the standard deviation (STDDEV) of (a) temperature, (b) zonal wind, and (c) meridional wind between the experiment (AMSU-A) run and the control (CNTL) run, derived against the ERA5 reanalysis. Hatched colors indicate the latitude regions (global; grey, Northern Hemisphere; blue, tropics; green, and Southern Hemisphere; red). Horizontal lines indicate 99% confidence intervals.