Dear Topical editor,

Thank you for your comment. As Dr. Juan A. Añel pointed out, the data availability statement is updated in the revised manuscript as below. As well as the scripts used for the data analysis and visualization, the source codes of the models (i.e., DART and CESM) used in this study are uploaded at https://doi.org/10.5281/zenodo.7714755. In addition, the model outputs are saved at https://doi.org/10.5281/zenodo.7983459, including the DART-derived analysis and the model 6-h forecasts as the model backgrounds.

[New]

Code and data availability.

DART version 9.11.13 was obtained from https://github.com/NCAR/DART. CESM version 2.1.0 is released at https://github.com/ESCOMP/CESM/tree/release-cesm2.1.0. Atmospheric initial conditions and the baseline observations at the BUFR format were obtained from the NCAR RDA (https://rda.ucar.edu/datasets/ds337.0 or https://doi.org/10.5065/Z83F-N512). AMSU-A Level-1B version 5 data from the Aqua satellite, including the calibrated brightness temperatures, were downloaded from the NASA Goddard Earth Sciences Data and Information Services Center (https://www.earthdata.nasa.gov/eosdis/daacs/gesdisc). In addition, AMSU-A Level-1B from NOAA-19, MetOp-A, and MetOp-B satellites were downloaded from the atmosphere product section in the EUMETSAT product navigator (https://navigator.eumetsat.int). The ECMWF ERA5 hourly data on pressure levels were acquired from the Copernicus Climate Change Service (C3S) Climate Data Store (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels). As well as the software codes, the model outputs are available at https://doi.org/10.5281/zenodo.7714755 and https://doi.org/10.5281/zenodo.7983459.

Referee #1

*Thank you very much for your substantial contribution. I like your study and how it is presented. Please clarify one issue: In figure 8, 10 and 11 you write "standard deviation (STDDEV) of the 500 hPa geopotential height" which could mean ensemble spread, but actually you mean "standard deviation of first-guess departures of the 500 hPa geopotential height". Please indicate in the caption somehow that it is the standard deviation of first-guess departures and *not* the standard deviation of the forecast. For example, write "standard deviation of first-guess departures".*

[Reply]
Thank you for your comment. In Figures 8, 10, and 11, the errors (i.e., mean bias and standard deviation) of four atmospheric variables (i.e., 500 hPa GPH, temperature, and zonal/meridional wind) were computed by comparing the DART-derived analysis (not first guess) with the ERA5 reanalysis as the reference. To clarify this point, we revised these captions as follows:

[Old]
"Figure 8. (a) Mean bias and (b) standard deviation (STDDEV) of 500 hPa geopotential height over the global (grey), Northern Hemisphere (NH: blue), tropics (TR: green), and Southern Hemisphere (SH: red), derived against the ERA5 reanalysis. Filled and hatched bars indicate the results for the control (CNTL) and experiment (AMSU-A) run, respectively. The 99% confidence intervals are indicated by the vertical black lines."
[New]
"Figure 8. (a) Mean bias and (b) standard deviation (STDDEV) of the analysis of 500 hPa geopotential height over the global (grey), Northern Hemisphere (NH: blue), tropics (TR: green), and Southern Hemisphere (SH: red), derived against the ERA5 reanalysis. Filled and hatched bars indicate the results for the control (CNTL) and experiment (AMSU-A) run, respectively. The 99% confidence intervals are indicated by the vertical black lines."

[Old]
"Figure 10. Spatial distribution of the standard deviation (STDDEV) of the 500 hPa geopotential height for the (a) control run (CNTL) and (b) experiment (AMSU-A) runs, derived against the ERA5 reanalysis."
[New]
"Figure 10. Spatial distribution of the standard deviation (STDDEV) of the analysis of 500 hPa geopotential height for the (a) control run (CNTL) and (b) experiment (AMSU-A) runs, derived against the ERA5 reanalysis."

[Old]
"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."
[New]
"Figure 11. Normalized difference of the standard deviation (STDDEV) of the analysis 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.”

Referee #2

Thank the authors to address my previous comments. The manuscript has been significantly improved. I have only one comment as below.
Several pre-trial runs were performed in the bias correction, error estimation, and etc. Could the authors please provide some details about the experiment designs and model settings of the pre-trial runs?

[Reply]
Thank you for your comment. As you mentioned, many pre-trial runs were performed to estimate the bias correction coefficients (i.e., scan-bias correction coefficients), optimal spatial thinning, localization half-width, and the observation errors. As the pre-trial runs were run to determine the specific values of each assimilation factor, the detailed setup was different for each pre-trial run depending on which factor was determined. We revised some sentences as follow:

[Old, lines 221-222]
"To choose the optimal spatial thinning distance, we performed four extra assimilation runs in which different spatial thinning distance (i.e., 96 km, 192 km, 288 km and 384 km) was applied."
[New]
"To choose the optimal spatial thinning distance, we performed four extra assimilation runs in which different spatial thinning distance (i.e., 96 km, 192 km, 288 km and 384 km) was applied. Except for the spatial thinning distance, these pre-trial runs were set up with the same assimilation factors, i.e., the estimated bias correction coefficients (refer to section 4.3), the estimated observation errors (refer to section 5), and the localization half-width of 0.075 (refer to section 6)."

[Old, lines 252-254]
"First, the mean bias of the departure between the AMSU-A observed radiances and forward-modeled radiances for each FOV is made with the data assimilation results derived from the pre-trial run."
[New]
"First, the mean bias of the departure between the AMSU-A observed radiances and forward-modeled radiances for each FOV is made with the data assimilation results derived from the pre-trial run. The pre-trial run was set up with the spatial thinning of 96 km (refer to section 4.2) and the default localization half-width (0.15, refer to section 6). The instrument noise errors were used as the observation errors within DART."

[Old, lines 324-325]
"In the pre-trial run, the instrument noise errors were initially used as the observation errors within DART."
[New]
"In the pre-trial run, instrument noise errors were simply used as the observation errors. The pre-trial run was set up with the default localization half-width (0.15, refer to section 6), the spatial thinning of 96 km (refer to section 4.2), and the bias correction scheme (refer to section 4.3)."

[Old, lines 360-361]
"To determine the localization half-width, three extra assimilation experiments were run with different half-widths (i.e., 0.15, 0.075, and 0.0375)."
[New]

“To determine the localization half-width, three extra assimilation experiments were run with different half-widths (i.e., 0.15, 0.075, and 0.0375). Except for the localization half-width, the assimilation experiments were set up with the spatial thinning of 96 km (refer to section 4.2), the bias correction scheme (refer to section 4.3), and the estimated observation errors (refer to section 5).”