Dear Prof. John Plane, Dear Jan Laštovička, reviewer#2, and Prof. Martin Mlynczak:

Thanks very much for taking your time to review our manuscript "*Trends of the high latitude mesosphere temperature and mesopause revealed by SABER (ID: egusphere-2024-396)*". We thank the reviewers for the time, insight, and effort that they have put into reviewing our manuscript. Those comments are all valuable and very helpful for revising and improving our paper.

Accordingly, we have uploaded a copy of the original manuscript with all the changes highlighted by using the track changes mode in MS Word. Appended to this letter is our point-by-point response to the comments raised by the reviewers. The original comments by reviewers use black, and our response is located below the comments and uses blue font.

Yours sincerely,

Xiao Liu, Jiyao Xu, Jia Yue, Yangkun Liu, and Vania F. Andrioli

----

Responses to the comments from Prof. Jan Laštovička (Reviewer#1)

The paper deals with SABER-derived long-term trends of mesospheric temperatures. This is very comprehensive and carefully done analysis. Authors use real (not constant) Yaw cycle lengths. They are making corrections of several factors, which can affect resulted trends, including reduction of seasonal variation effect using climatological NRLMSISE2.1 mean temperature. So many corrections may introduce some imperfection (e.g. couple of years ago NRLMSISE00 was used, now we are using version 2.1 and after some years again a new and more accurate version appears). However, authors made best corrections at the level of present-day knowledge. I require only very minor essentially technical corrections, so after making these corrections I recommend the paper to publication.

Response: We really appreciate your efforts in reviewing our manuscript and your clear and detailed feedback. Following your comments and suggestions, we have revised the manuscript accordingly. Please find the point-to-point responses below.

Comments:

1. Line 370: Cooling trend is becoming larger with increasing height, so "decrease" should be replaced by "increase".

   Response: Thanks for your suggestion. We have revised these descriptions throughout the manuscript. If cooling trends become more negative (i.e., having large absolute value), we state that "cooling trends increase".

Wording and misprints:

1. Line 33: delete "and"

   Response: It has been deleted. In the new version, this sentence is revised as "The corrected temperature without any waves is regarded as the mean temperature".

2. Line 71: "of a cooling trend" should be "cooling trends"

   Response: Following your suggestion, we have revised this sentence as "Garcia et al. (2019) showed that the global mean SABER temperature (52°S–52°N) cooling trends…".

3. Line 72: "mesosphere, were" should be "mesosphere were"

   Response: This sentence has been broken into two short sentences. The revised version is "…Garcia et al. (2019) showed that the global mean SABER temperature (52°S–52°N) had cooling trends of 0.4–0.5 K/decade during 2002–2018 in the stratosphere and mesosphere. These magnitudes were smaller than those simulated by WACCM4 (0.6–0.9 K/decade) but within 2 times of the standard deviation".

4. Line 323: "7(b)" should be "5(b)"

   Response: We are sorry for the mistake. Now, it has been revised as "Figures 5(a) and 5(b) show the mean $\overline{T}_{msp}$ and z_{msp} over 22 years…".

5. Lines 403 and 480: "The peaks" should be better "The heights of peaks"

   Response: Following your suggestion and the comments from reviewer#2. Now, it has been revised as "The altitudes of peaks of the warming trends vary from 4×10^{-3} hPa to 10^{-3} hPa in different months".

6. Line 464: "excludling" should be "excluding".

   Response: Following your suggestion. Now, it has been revised as "Then we can obtain an accurate mean temperature excluding atmospheric waves".

Responses to the comments from Reviewer#2

The paper highlights the MLT temperature trend in high latitudes through a new innovated SABER data processing algorithm that solves the previous issue (fixed 60-day window) regarding the forward drifting of SABER local time coverage, while mitigating properly the embedded bias due to seasonal variations through the assistant of MSIS2.0. The results show mostly cooling trends

around the globe with sporadic spots of warming, which is consistent with the numerical studies. The author states that the revealed large cooling at high latitudes MLT demonstrates the high sensitivity to the global climate change in this area. Note that similar statement has been raised by the climate studies focusing on the troposphere and stratosphere. The manuscript and figures are mostly clean, and I just have a few minor questions that need the author to address.

Response: Thank you very much for your time involved in reviewing the manuscript and your very encouraging comments. Please find the point-to-point responses below.

Comments:

1. Based on equation 1, removing kt_{UT} from the mean $\bar{T}_d$ should give mean $\bar{T}_{d0}$, instead of the residual term. Please clarify.

   Response: Thanks for your careful reading. In the new version, this point has been clarified as:

   $$\bar{T}_d = \bar{T}_{d0} + kt_{UT} + \bar{T}_{res}. \tag{1}$$

   Here, $\bar{T}_{d0}$ is the mean temperature in each YC. t_{UT} is the universal time with a unit of day, k represents the linear variation of $\bar{T}_d$ in each YC. After removing $\bar{T}_{d0}$ and the linear variation (kt_{UT}) from $\bar{T}_d$, we get a residual temperature $\bar{T}_{res}$ of each YC.

2. Some of the trend profiles at high altitudes (geometric height) in figure 6 showing near or more than 10 K/decade near 100 km and above, even considering the fitting uncertainty. I feel these cooling trends are a little excessive. The author might want to double check the data quality or the algorithm for this altitude range.

   Response: Following your comment, we discussed the uncertainties of the SABER temperature and their possible impacts on the extreme cooling trend. To make the discussions clearly, we rearrange the section of Discussions (Section 4) as three subsections:

   Sec. 4.1 for "The reliability of trends in the MLT region at latitudes lower than 50°N/S";

   Sec. 4.2 for "The reliability of trends in the MLT region at latitudes higher than 50°N/S";

   Sec. 4.3 for "The reliability of the mesopause trends".

   The Sec. 4.1 and 4.3 do not change much. The main revisions are included in Sec. 4.2 and in the Appendix. In Sec. 4.2, the uncertainties of SABER temperature measurements on the derived trends are discussed on the following three aspects: (1) the SABER measurement errors for temperature; (2) the drivers' uncertainties in retrieving the SABER temperature; (3) the impacts of the measurement uncertainties on the derived trends.

   The detailed revisions in the text are provided below:

   (1) The description on the SABER measurement errors for temperature (in the Introduction)

The operational SABER temperature profile covers an altitude range of ~15–110 km. The uncertainties of SABER temperature profile are height dependent. For a single temperature profile, its uncertainties are summarized at https://spdf.gsfc.nasa.gov/pub/data/timed/saber/ and are of ~1.8–2.3 K at z=60–80 km, ~5.4–8.4 K at 90–100 km, and ~8.4–29.2 K at 100–110 km under the condition of vertical resolution of 2 km (Remsberg et al., 2008; Rezac et al., 2015; Dawkins et al., 2018)".

Dawkins, E. C. M., Feofilov, A., Rezac, L., Kutepov, A. A., Janches, D., Höffner, J., Chu, X., Lu, X., Mlynczak, M. G., and Russell, J.: Validation of SABER v2.0 operational temperature data with ground-based lidars in the mesosphere-lower thermosphere region (75–105 km), J. Geophys. Res. Atmos., 123, 9916–9934, https://doi.org/10.1029/2018JD028742, 2018.

(2) The discussions on the drivers' uncertainties in retrieving the SABER temperature (in Sec. 4.2)

The main causes of the operational SABER temperature systematic uncertainties are the lack of accurate knowledge of atomic oxygen and carbon dioxide during the retrieval process. The atomic oxygen provided to the operational SABER temperature retrieval algorithm is from NRLMSISE-00 (Picone et al., 2002). Below 100 km, no atmospheric observations of atomic oxygen are incorporated. Thus, the uncertainty of atomic oxygen influences the uncertainties of temperature from ~75 km to 110 km, in particular, above 100 km. The carbon dioxide provided to the operational SABER temperature retrieval algorithm is the monthly average value from WACCM (Dawkins et al., 2018; Picone et al., 2002). Thus, there is no local time variation in carbon dioxide used in the operational SABER temperature retrieval algorithm. This will induce uncertainties of SABER temperature and thus the uncertainties of trends above 75 km.

(3) The discussions on the impacts of the measurement uncertainties on the derived trends through Monte Carlo simulation (in Sec. 4.2 and Appendix):

In Sec 4.2, the followings have been included:

These uncertainties in temperature may not be constant or stable in time or in space. To explore the impacts of the uncertainties in SABER temperature on the derived trends, we performed Monte Carlo simulations by assuming the uncertainties in SABER temperature following a uniform distribution in the range of ±25K. In each time of Monte Carlo simulation, in each YC and at each pressure level and within a latitude band of 10°, the SABER samplings (more than 5000 data) are added by random numbers following the uniform distribution in the range of ±25K. Then same procedure described in Sec. 2.1–2.3 was repeated to derive trends. The Monte Carlo simulations were performed 5000 times (see Appendix). The main result is that the uncertainties of ±25K in SABER

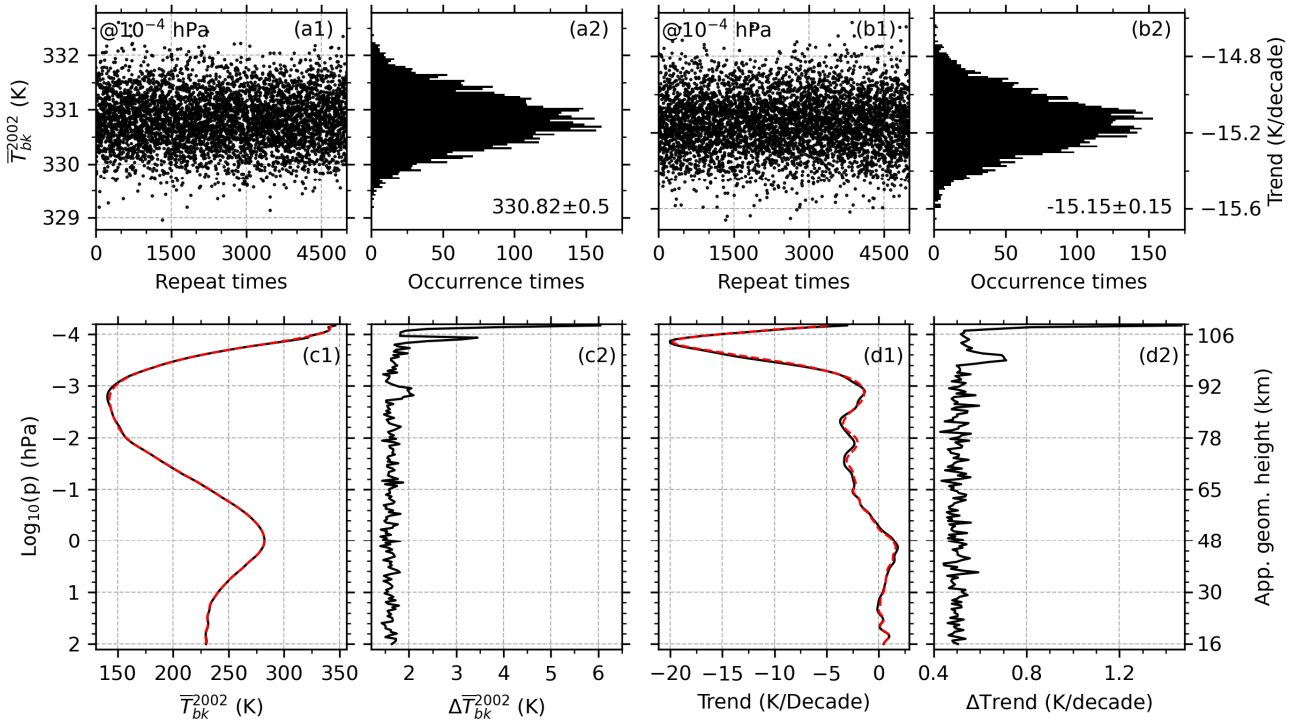samplings would induce a mean temperature variation of ~1–3 K and a false trend ~0.5–1.2 K/decade at high latitudes. This is mainly because mean temperature is calculated from more than 5000 data in each YC within a latitude band of 10°, which reduces the standard deviation by a factor of ~1/250 based on central limit theory. It must be noted that the actual distributions of the uncertainties in SABER samplings caused by atomic oxygen and carbon dioxide are unknown. The Monte Carlo simulation only provides a reference result by assuming the uncertainties following uniform distributions. This may not be valid for the case of SABER temperature systematic errors. So may not be valid. We only include it in the Appendix.

In the Appendix, the followings have been included:

Around 10^{-4} hPa, the uncertainties of SABER temperature measurements are around 25 K at mid-latitudes and are likely higher at high latitudes. These uncertainties are mainly attributed to the uncertainties of atomic oxygen and carbon dioxide, which were used in the operational SABER temperature retrieval algorithm. Moreover, these uncertainties in temperature may not be constant or stable in time or in space. To explore the impacts of the uncertainties in SABER temperature on the derived trends, we performed Monte Carlo simulations by assuming the uncertainties in SABER temperature following a uniform distribution in the range of ±25K. In each time of Monte Carlo simulation, in each YC and at each pressure level and within a latitude band of 10°, the SABER samplings (more than 5000 data) are added by random numbers which follows the uniform distribution in the range of ±25K. Then same procedure described in Sec. 2.1–2.3 was repeated to derive trends. The Monte Carlo simulations were performed 5000 times to get convincing results.

Since the cooling trends are very large in YC3 and at 75°N, especially around the pressure levels of around 10^{-4} hPa, we show in Figure A the impact of the random uncertainties of SABER temperature on the derived trends in YC3 and at 75°N. The uncertainties of ±25K in SABER samplings induce the mean temperature ($\overline{T}_{bk}^{2002}$) varying in the range of ±2 K (Fig. Aa1) with standard deviation of 0.5 K (Fig. Aa2) at 10^{-4} hPa. This in turn induces the trends varying in the range of ±0.6 K/decade (Fig. Ab1) with standard deviation of 0.15 K/decade (Fig. Ab2) at 10^{-4} hPa. The altitude profile of $\overline{T}_{bk}^{2002}$ by assuming a zero uncertainty is similar to that calculated by assuming the uncertainties of ±25K (Fig. Ac1). The differences of the maximum and minimum of $\overline{T}_{bk}^{2002}$ among the 5000 times of Monte Carlo simulations are ~1–2 K below 5×10^{-4} hPa and are ≥3 K around 10^{-4} hPa (Fig. Ac2). The altitude profile of trend by assuming a zero uncertainty is similar to that calculated by assuming the uncertainties of ±25K (Fig. Ad1). The differences of the maximum and minimum of trends among the 5000 times of Monte Carlo simulations are ~0.5 K/decade below 10^{-3} hPa and are ~0.5–1.2 K/decade around 10^{-4} hPa (Fig. Ad2). This example illustrates that the uncertainties of ±25K in SABER samplings would induce a mean temperature variation of ~1–3 K

and a false trend ~0.5–1.2 K/decade at high latitudes.

Figure A. The impacts of uncertainties of ±25K in SABER temperature on the derived trends in YC3 and at 75°N during 5000 times of Monte Carlo simulation. (a1) and (a2): the mean temperature calculated from SABER sampling ($\overline{T}_{bk}^{2002}$) and its histogram at 10^{-4} hPa; (b1) and (b2): the trend and its histogram at 10^{-4} hPa; (c1) and (d1): the altitude profiles of $\overline{T}_{bk}^{2002}$ by assuming zero uncertainty (black) and uncertainties of ±25K (dashed-black); (c2) and (d2) altitude profile of the difference between the maximum and minimum of $\overline{T}_{bk}^{2002}$ and trend.

Another Monto Carlo simulation is performed to test the impacts of the uncertainties of ±25K on the mean temperature (180 K) by changing the sampling points. During 5000 times of simulations (Figures R1 and R2, which are only shown here but not in the text), the mean temperature and its standard deviation are 179.956±4.5 K if there are 10 samplings; the mean temperature and its standard deviation are 179.977±1.43 K if there are 100 samplings; the mean temperature and its standard deviation are 179.997±0.20 K if there are 5000 samplings. This indicates that the increasing samplings can reduce the measurement uncertainties efficiently. Although the uncertainties of SABER samplings are as large as ±25K at high latitudes, its impact on the trends are insignificant in the highly averaged results. This is mainly because mean temperature is calculated from more than 5000 data in each YC within a latitude band of 10°, which reduces the standard deviation by a factor of ~1/250 based on central limit theory. It must be noted that the actual distributions of the uncertainties in SABER samplings are unknown. The Monte Carlo simulation only provides a

reference result by assuming the uncertainties following uniform distribution. This may not be valid for the case of SABER temperature systematic errors. We only include it in the Appendix.

The followings are the description of the central limit theory (CLT) but not included in the text:

The mathematic basis of the highly averaged result has very small standard deviations is the central limit theory (CLT) in probability and statistics. Suppose random variables $\{X_i\}_{i=1}^{n}$ are independent and identically distributed and have an expectation of μ and standard deviation of σ, the distribution function,
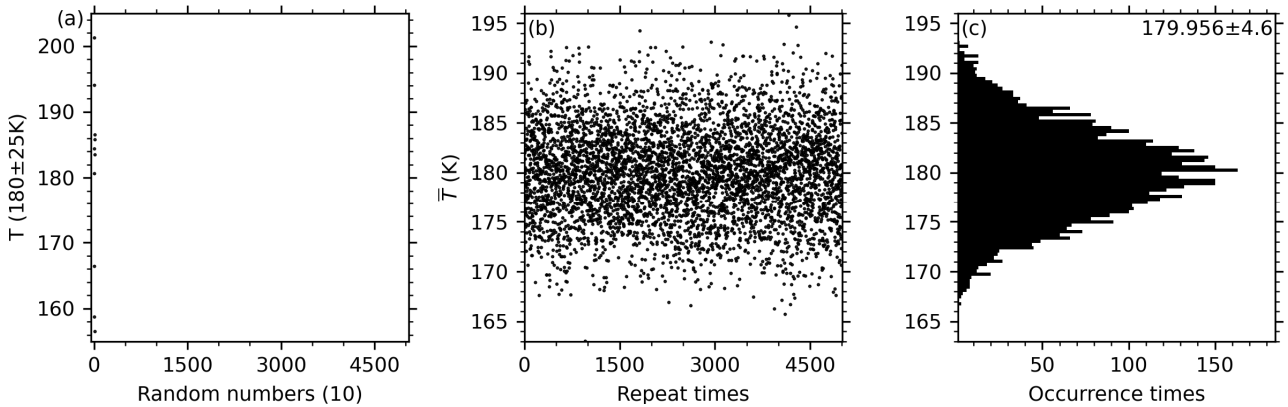
$$F_n(x) = P\left\{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \le x\right\},$$

has limitation of,
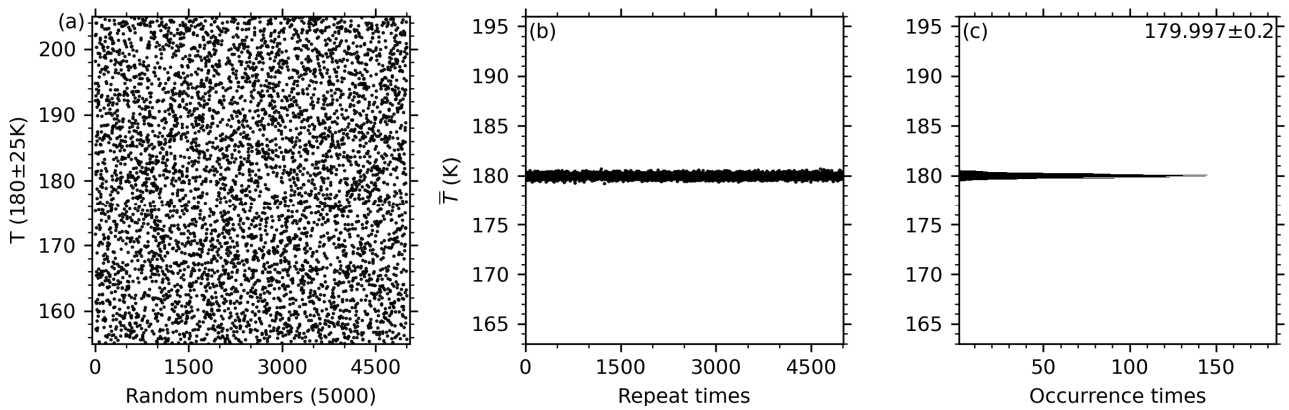
$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P\left\{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \le x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

The CLT states that if one take sufficiently large samples from a population, the samples' means will be normally distributed, even if the population isn't normally distributed. Thus, X follows the normal distribution of $N\left(\mu, \sigma/\sqrt{n}\right)$.

The uniform distribution in the range of $[a, b]$, its expectation and standard deviation are $\mu = (a + b)/2$ and $\sigma = (b - a)/\sqrt{12}$, respectively. According to CLT, the uncertainties of $\pm 25K$ will induce an uncertainty of $50/\sqrt{12 \times 5000} \approx 50/245 = 0.204K$. This support the Monte Carlo simulations of ± 0.20 K if there are 5000 samplings.

Figure R1. *Monte Carlo simulation on the influences of 10 samplings on the mean and standard deviation of the uniform distribution of $\pm 25K$ with mean of 180 K. (a) an example of the 10 samplings; (b) and (c) show, respectively, the means and their histogram during 5000 times Monte Carlo simulations. The mean and standard deviation are labelled on the top right corner of (c).*

Figure R2. *Same caption as Fig. R1 but for 5000 samplings.*

3. Line 42, "highest latitudes" why not just say high latitudes?

   Response: Following your suggestion, the "highest latitudes" is revised as "high latitudes".

4. Line 95, "lower heights", please be more specific, troposphere or stratosphere? Or just say lower altitudes.

   Response: The CO_2 data used in the LIMA model were measured at Mauna Loa (19°N, 155°W) and were considered according to observations in the troposphere (Lübken et al., 2021).

   In the new version, we have specified "lower heights" as "troposphere". The revised sentence is "…the negative trend of noctilucent clouds altitudes (~83 km) was primarily caused by the increasing CO_2 in the troposphere during…".

5. Line 121, 60-day

   Response: It is revised in the new version.

6. Figure 4 caption. I do not see red and green dots, but purple and blue ones. Also, it is very difficult to tell the "+" signs. May want to change the symbol.

   Response: You are right. This sentence is revised as "The purple and blue dots indicate the heights of the mesopause and stratopause, respectively". The "+" signals are too small to readable. We have changed them as shaded points. The revised figure and its caption are shown below and are also included in the text.
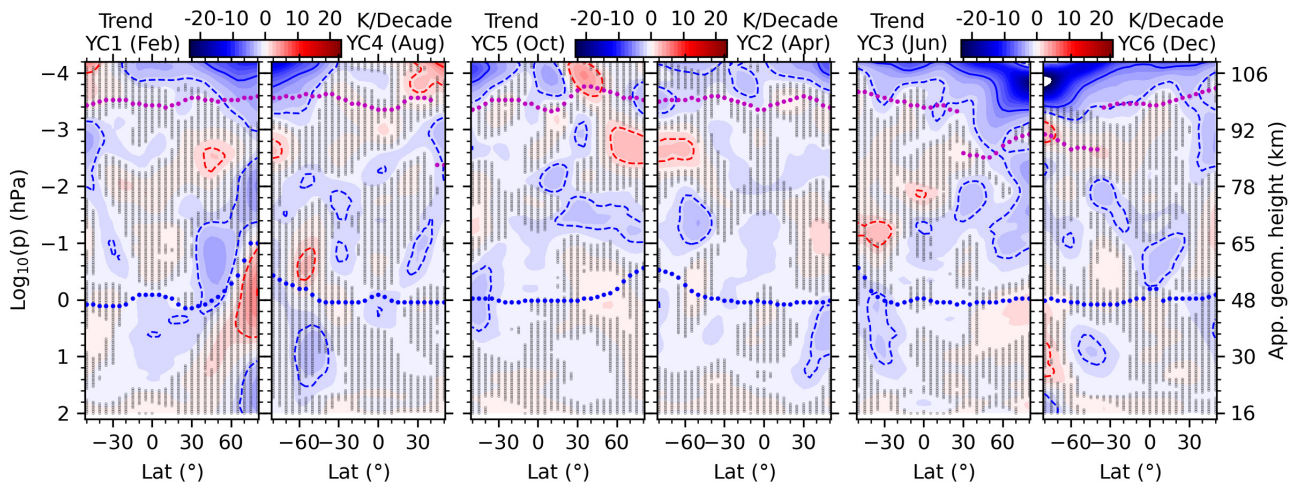
Figure 4. Trends of the corrected mean temperature in the six YCs. The solid and dashed contour lines indicate ±6 and ±2 K/decade, respectively. The purple and blue dots indicate the heights of the mesopause and stratopause, respectively. The regions marked by shaded points indicate that trends are not significant with reference to the 95% the confidence level. The approximate geometric height is label on the last panel.

7. Line 370, "higher heights", again please be specific. And replace all the "heights" with "altitudes".

    Response: Here, the "higher heights" is specified as "Above 5×10^{-3} hPa". Following your suggestion, we replaced replace all the "heights" with "altitudes" in the new version.

Responses to the comments from Prof. Martin Mlynczak (Reviewer#3)

Overview:

This paper presents a study of trends in temperature at high latitudes using data from the SABER instrument on the TIMED satellite. The trends are determined by a standard linear regression procedure. The paper clearly understands the issue with the 'moving' yaw cycle and presents analyses which attempt to account for that. While mostly in accord with other trend studies involving SABER data, the paper presents some very remarkable trend values (6 K/decade to 10 K/decade) which are well beyond what is expected if the trends are due solely to the radiative response to increasing greenhouse gases. The paper also appears to lack any consideration of measurement uncertainty of the SABER temperature parameter and the impact of these uncertainties on the uncertainty of the derived trends.

Response: We appreciate the time and effort you have taken to review our work. Your thoughtful and constructive comments helped us to improve the quality of our manuscript. Following your comments, we discussed the possible reasons including the measurement uncertainties of the SABER temperature data behind the unexpected large trends.

(1) For the unexpected trend values (6 K/decade to 10 K/decade) in the high latitude MLT region

The unexpected trend values in the high latitude MLT region might be a combination effects of both radiative (i.e., Garcia et al., 2019, Mlynczak et al., 2022) and dynamical feedback (Beig et al., 2003; Beig, 2011). The dynamical feedback has been discussed based on the simplified Eulerian mean thermodynamic equation and is supported by the increasing trends of gravity waves in the summer hemispheres.

(2) The measurement uncertainties of the SABER temperature and their impact on the uncertainty of the derived trends

Following your recommendation below, we discussed the measurement uncertainties based on the knowledge of atomic oxygen and carbon dioxide used in the SABER temperature retrieval algorithm. Moreover, we performed 5000 rounds of Monte Carlo simulation to illustrated the measurement uncertainties on the derived trends (in the Appendix). By assuming the uncertainties following unform distribution in the range of ±25 K in SABER samplings, the simulation results show that these uncertainties would induce a mean temperature variation of ~1–3 K and a false trend of ~0.5–1.2 K/decade at high latitudes. This neglectable influence is mainly because the mean temperature is calculated from more than 5000 data in each YC within a latitude band of 10°, which reduce the standard deviation by a factor of ~1/250 based on central limit theory. It must be noted

that the actual distributions of the systematic uncertainties in SABER samplings are unknown. The Monte Carlo simulation only provides a reference result by assuming the uncertainties following uniform distribution. This may not be valid for the case of SABER temperature systematic errors. So may not be valid. We only include it in the Appendix.

Recommendation:

There are a couple of major issues which the authors need to address and that relate to uncertainties/errors in the SABER temperature data. The authors must convincingly address these before the paper can be considered for publication.

Response: Following your clear recommendations in below, we discussed the uncertainties of the SABER temperature and their possible impacts on the extreme cooling trend. To make the discussions clearly, we rearrange the section of Discussions (Section 4) as three subsections:

Sec. 4.1 for "The reliability of trends in the MLT region at latitudes lower than 50°N/S";

Sec. 4.2 for "The reliability of trends in the MLT region at latitudes higher than 50°N/S";

Sec. 4.3 for "The reliability of the mesopause trends".

The Sec. 4.1 and 4.3 do not change much. The main revisions are included in Sec. 4.2 and in Appendix.

The major issues have been addressed on the following four aspects: (1) the ability of F10.7 in indicating the effects solar radiation on the lower thermosphere; (2) the dynamical feedback that causes additional cooling; (3) the warmer trends in the polar troposphere as compared to those at lower and middle latitudes; (4) the uncertainties of SABER temperature measurements and their impacts on the derived trends.

In the Appendix, we performed 5000 rounds of Monte Carlo simulation to explore the impacts of the uncertainties in SABER temperature on the derived trends.

Please see the point-to-point responses in below.

Comments:

1. The large trends identified in the polar region (ranging from 6 K/decade to 10 K/decade) are presented without discussion of possible effects of measurement error and without discussion of their physical meaning or likelihood. Mlynczak et al. (2022) noted that the expected global average mesosphere temperature change to a doubling of CO_2 (i.e., the climate sensitivity) was about 6.5 K. The paper is presenting results that imply a climate sensitivity of the polar mesosphere of about 10 times that. What would be the physical mechanism for a solely radiative effect that would make the polar mesosphere 10 times more sensitive to CO_2 increase than the global average? Is there a radiative or dynamical feedback that causes additional cooling besides

what might be expected from a purely radiative effect? It is important to understand this point because a 10 K/decade trend would result in non-physical temperatures in a few decades and would also imply a substantially hotter polar upper mesosphere and lower thermosphere at the start of the Industrial Age. We are about halfway to doubled CO_2 now and so addressing this issue is critical to placing the results and their consequences in perspective.

Response: Besides the radiative cooling caused by CO_2, the derived cooling trends might be caused by (1) the ability of F10.7 in representing the variation of solar radiation on the lower thermosphere, (2) the dynamical feedback in the polar MLT region,

These possible reasons have been included in text (Sec. 4.2):

We can see the extreme cooling trends of ≥6 K/decade above ~10^{-3} hPa in YC3 and YC6 and in YC1 and YC4 but around 10^{-4} hPa. These cooling trends are comparable with the global average mesosphere temperature of 6.8–8.4 K/decade derived by Mlynczak et al. (2022) after doubling of CO_2 at Earth's surface. However, It takes decades to doubled CO_2. Thus, a purely radiative effect due to the increasing CO_2 cannot support the extreme cooling trends derived here. Mlynczak et al. (2022) proposed that the F10.7 is not a suitable proxy to indicate effects of the solar radiations on the lower thermosphere. But the solar irradiance in the Schumann–Runge band (175–200 nm) might be responsible for the colder trend. Even so, the extreme cooling trends of ~10 K/decade are still larger than those reported by Mlynczak et al. (2022). Other possible reasons for the extreme cooling trends in the high latitude MLT region can be attributed to: (1) the dynamical feedback in the polar MLT region; (2) the uncertainties of the SABER temperature measurements.

Besides the purely radiative effect on the cooling trends in the MLT region (i.e., Garcia et al., 2019, Mlynczak et al., 2022), the dynamical feedback might be another cause of the cooling trends. Based on the simplified Eulerian mean (TEM) thermodynamic equation, the temperature change (ΔT) caused by dynamics can be written as (Eq. 3 and 4 of Yu et al. (2023)),

$$\Delta T = -\alpha^{-1}\left(w^* S + v^* \frac{\partial \bar{T}}{a\partial \varphi}\right). \tag{4}$$

Here, α is the Newtonian cooling coefficient. w^* and v^* are the residual vertical and meridional velocity, respectively. S and $\bar{T}$ are the static stability and zonal mean temperature, respectively. a and φ are the Earth's radius and latitude, respectively. From Eq. (4), we propose that the extreme cooling trends at high latitudes of the summer hemispheres (YC3 and YC6) might be resulted from the changing summer-to-winter circulation and gravity wave forcing in the MLT region. The circulation is upwelling (positive w^*) in the summer hemisphere and causes a cold summer mesosphere through adiabatic cooling. Conversely, in the winter hemisphere, the circulation is downwelling (negative w^*), leading to a warm winter mesosphere through adiabatic warming (Garcia and Solomon, 1985). A necessary condition for the extreme cooling trends at summer high

latitudes is the stronger upwelling and thus the increasing gravity wave body force in the summer hemispheres. Previous studies showed that the potential energy of gravity waves (GWPE) in the MLT region exhibited significant positive trends at southern high latitudes in January and at northern high latitudes in July (Fig. 5 of Liu et al., 2017). The positive trends of GWPE might enhance the strength of upwelling and thus result in the extreme cooling trends at high latitudes of summer hemispheres. It should be noted that the dynamical feedback in the MLT region is only analyzed qualitatively, the quantitative analysis should be performed through model simulations. Such that one can elucidate the physics behind the strong cooling trend in the polar MLT region.

2. In order to believe the large, derived trends, all analyses must consider the uncertainty in the SABER temperature data, particularly in polar regions, and particularly at the lowest pressure levels (highest altitudes). The paper cites papers by Remsberg and Rezac in temperature uncertainties below 100 km. The Rezac paper is for a version of the SABER data that is not used by the authors. The authors are referred to this link for a summary of SABER measurement errors for temperature: https://saber.gats-inc.com/temp_errors.php

In particular, the paper states there is a trend of 10 K/decade (line 296) at 10^{-4} hPa. However, the uncertainty at this pressure level is 25 K at mid-latitudes and it is likely higher in polar regions. The main drivers of SABER temperature uncertainty are the knowledge of atomic oxygen and carbon dioxide which are provided to the SABER temperature algorithm by the MSIS 2000 model and by the WACCM model, respectively.

The MSIS 2000 model is over 20 years old and has incorrect local time variations in atomic oxygen as has been noted in the literature. In addition, below 100 km, no atmospheric observations of atomic oxygen are incorporated into the MSIS 2000 model. It must be assumed that the atomic oxygen (which influences the uncertainty on temperature from ~ 75 km to 110 km) is uncertain in the polar regions and there are corresponding uncertainties in temperature.

Furthermore, monthly average values of CO_2 used in the derivation of temperature are provided by the WACCM model. There is no local time variation in CO_2 used in the SABER retrieval. SABER temperatures, particularly above 80 km, are very sensitive to the CO_2 abundance.

In essence, for the trends in temperature to be correct, the variability and trends in O and CO_2 provided by MSIS 2000 and WACCM must also be correct. There is no real way to validate if this is true in the polar regions.

As noted above, the uncertainty of SABER data at mid latitudes is 25 K at 10^{-4} hPa. It may be

higher in the polar regions during summer due to the low temperatures. The key point is that the uncertainty at any altitude does not necessarily cancel out when computing trends because the error in temperature due to O and to CO_2 may not be constant or even the same sign over time. This may be thought of as a mild form of algorithm instability in which the inputs to the temperature algorithm do not represent the actual atmosphere and consequently cause uncertainty on the retrieved temperature. The uncertainty in temperature may not be constant in time.

The recommendation to the authors is to compute the uncertainty in the trend assuming the errors on the temperatures are non-zero and follow standard error analyses for uncertainty calculations when taking differences. At what point do the uncertainties in temperature negate the large trend values?

Response: Following your suggestions, the improvements are ascribed as the following three aspects:

(1) The description on the SABER measurement errors for temperature (in the Introduction)

The operational SABER temperature profile covers an altitude range of ~15–110 km. The uncertainties of SABER temperature profile are height dependent. For a single temperature profile, its uncertainties are summarized at https://spdf.gsfc.nasa.gov/pub/data/timed/saber/ and are of ~1.8–2.3 K at z=60–80 km, ~5.4–8.4 K at 90–100 km, and ~8.4–29.2 K at 100–110 km under the condition of vertical resolution of 2 km (Remsberg et al., 2008; Rezac et al., 2015; Dawkins et al., 2018)".

Dawkins, E. C. M., Feofilov, A., Rezac, L., Kutepov, A. A., Janches, D., Höffner, J., Chu, X., Lu, X., Mlynczak, M. G., and Russell, J.: Validation of SABER v2.0 operational temperature data with ground-based lidars in the mesosphere-lower thermosphere region (75–105 km), J. Geophys. Res. Atmos., 123, 9916–9934, https://doi.org/10.1029/2018JD028742, 2018.

(2) The discussions on the drivers' uncertainties in retrieving the SABER temperature (in Sec. 4.2)

The main causes of the operational SABER temperature systematic uncertainties are the lack of accurate knowledge of atomic oxygen and carbon dioxide during the retrieval process. The atomic oxygen provided to the operational SABER temperature retrieval algorithm is from NRLMSISE-00 (Picone et al., 2002). Below 100 km, no atmospheric observations of atomic oxygen are incorporated. Thus, the uncertainty of atomic oxygen influences the uncertainties of temperature from ~75 km to 110 km, in particular, above 100 km. The carbon dioxide provided to the operational SABER temperature retrieval algorithm is the monthly average value from WACCM (Dawkins et al., 2018; Picone et al., 2002). Thus, there is no local time variation in carbon dioxide used in the operational

SABER temperature retrieval algorithm. This will induce uncertainties of SABER temperature and thus the uncertainties of trends above 75 km.

(3) The discussions on the impacts of the measurement uncertainties on the derived trends through Monte Carlo simulation (in Sec. 4.2 and Appendix):
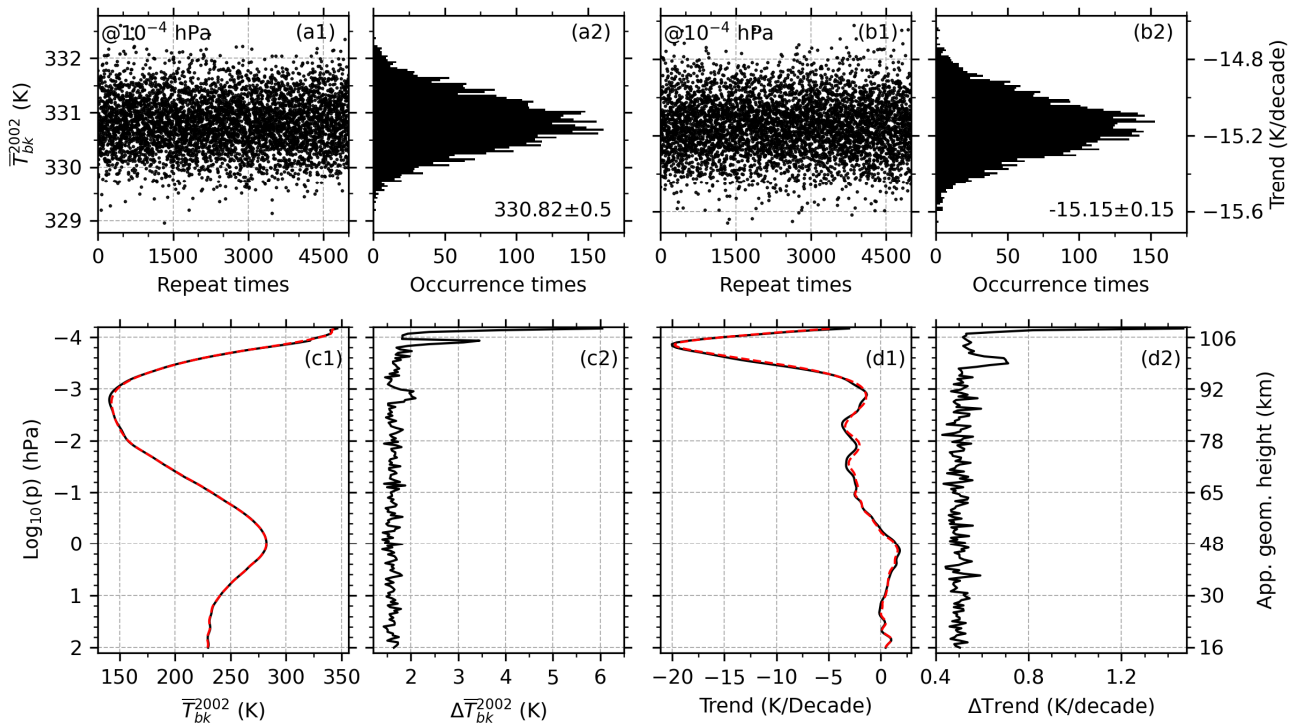
In Sec 4.2, the followings have been included:

These uncertainties in temperature may not be constant or stable in time or in space. To explore the impacts of the uncertainties in SABER temperature on the derived trends, we performed Monte Carlo simulations by assuming the uncertainties in SABER temperature following a uniform distribution in the range of ±25K. In each time of Monte Carlo simulation, in each YC and at each pressure level and within a latitude band of 10°, the SABER samplings (more than 5000 data) are added by random numbers following the uniform distribution in the range of ±25K. Then same procedure described in Sec. 2.1–2.3 was repeated to derive trends. The Monte Carlo simulations were performed 5000 times (see Appendix). The main result is that the uncertainties of ±25K in SABER samplings would induce a mean temperature variation of ~1–3 K and a false trend ~0.5–1.2 K/decade at high latitudes. This is mainly because mean temperature is calculated from more than 5000 data in each YC within a latitude band of 10°, which reduces the standard deviation by a factor of ~1/250 based on central limit theory. It must be noted that the actual distributions of the uncertainties in SABER samplings caused by atomic oxygen and carbon dioxide are unknown. The Monte Carlo simulation only provides a reference result by assuming the uncertainties following uniform distributions. This may not be valid for the case of SABER temperature systematic errors. So may not be valid. We only include it in the Appendix.

Around 10^{-4} hPa, the uncertainties of SABER temperature measurements are around 25 K at mid-latitudes and are likely higher at high latitudes. These uncertainties are mainly attributed to the uncertainties of atomic oxygen and carbon dioxide, which were used in the operational SABER temperature retrieval algorithm. Moreover, these uncertainties in temperature may not be constant or stable in time or in space. To explore the impacts of the uncertainties in SABER temperature on the derived trends, we performed Monte Carlo simulations by assuming the uncertainties in SABER temperature following a uniform distribution in the range of ±25K. In each time of Monte Carlo simulation, in each YC and at each pressure level and within a latitude band of 10°, the SABER samplings (more than 5000 data) are added by random numbers which follows the uniform distribution in the range of ±25K. Then same procedure described in Sec. 2.1–2.3 was repeated to derive trends. The Monte Carlo simulations were performed 5000 times to get convincing results.

Since the cooling trends are very large in YC3 and at 75°N, especially around the pressure levels

of around 10^{-4} hPa, we show in Figure A the impact of the random uncertainties of SABER temperature on the derived trends in YC3 and at 75°N. The uncertainties of ±25K in SABER samplings induce the mean temperature ($\overline{T}_{bk}^{2002}$) varying in the range of ±2 K (Fig. Aa1) with standard deviation of 0.5 K (Fig. Aa2) at 10^{-4} hPa. This in turn induces the trends varying in the range of ±0.6 K/decade (Fig. Ab1) with standard deviation of 0.15 K/decade (Fig. Ab2) at 10^{-4} hPa. The altitude profile of $\overline{T}_{bk}^{2002}$ by assuming a zero uncertainty is similar to that calculated by assuming the uncertainties of ±25K (Fig. Ac1). The differences of the maximum and minimum of $\overline{T}_{bk}^{2002}$ among the 5000 times of Monte Carlo simulations are ~1–2 K below 5×10^{-4} hPa and are ≥3 K around 10^{-4} hPa (Fig. Ac2). The altitude profile of trend by assuming a zero uncertainty is similar to that calculated by assuming the uncertainties of ±25K (Fig. Ad1). The differences of the maximum and minimum of trends among the 5000 times of Monte Carlo simulations are ~0.5 K/decade below 10^{-3} hPa and are ~0.5–1.2 K/decade around 10^{-4} hPa (Fig. Ad2). This example illustrates that the uncertainties of ±25K in SABER samplings would induce a mean temperature variation of ~1–3 K and a false trend ~0.5–1.2 K/decade at high latitudes.

Figure A. The impacts of uncertainties of ±25K in SABER temperature on the derived trends in YC3 and at 75°N during 5000 times of Monte Carlo simulation. (a1) and (a2): the mean temperature calculated from SABER sampling ($\overline{T}_{bk}^{2002}$) and its histogram at 10^{-4} hPa; (b1) and (b2): the trend and its histogram at 10^{-4} hPa; (c1) and (d1): the altitude profiles of $\overline{T}_{bk}^{2002}$ by assuming zero uncertainty (black) and uncertainties of ±25K (dashed-black); (c2) and (d2) altitude profile of the difference between the maximum and minimum of $\overline{T}_{bk}^{2002}$ and trend.

Another Monto Carlo simulation is performed to test the impacts of the uncertainties of ±25K on the mean temperature (180 K) by changing the sampling points. During 5000 times of simulations (Figures R1 and R2, which are only shown here but not in the text), the mean temperature and its standard deviation are 179.956±4.5 K if there are 10 samplings; the mean temperature and its standard deviation are 179.977±1.43 K if there are 100 samplings; the mean temperature and its standard deviation are 179.997±0.20 K if there are 5000 samplings. This indicates that the increasing samplings can reduce the measurement uncertainties efficiently. Although the uncertainties of SABER samplings are as large as ±25K at high latitudes, its impact on the trends are insignificant in the highly averaged results. This is mainly because mean temperature is calculated from more than 5000 data in each YC within a latitude band of 10°, which reduces the standard deviation by a factor of ~1/250 based on central limit theory. It must be noted that the actual distributions of the uncertainties in SABER samplings are unknown. The Monte Carlo simulation only provides a reference result by assuming the uncertainties following uniform distribution. This may not be valid for the case of SABER temperature systematic errors. We only include it in the Appendix.

The followings are the description of the central limit theory (CLT) but not included in the text:

The mathematic basis of the highly averaged result has very small standard deviations is the central limit theory (CLT) in probability and statistics. Suppose random variables $\{X_i\}_{i=1}^{n}$ are independent and identically distributed and have an expectation of μ and standard deviation of σ, the distribution function,
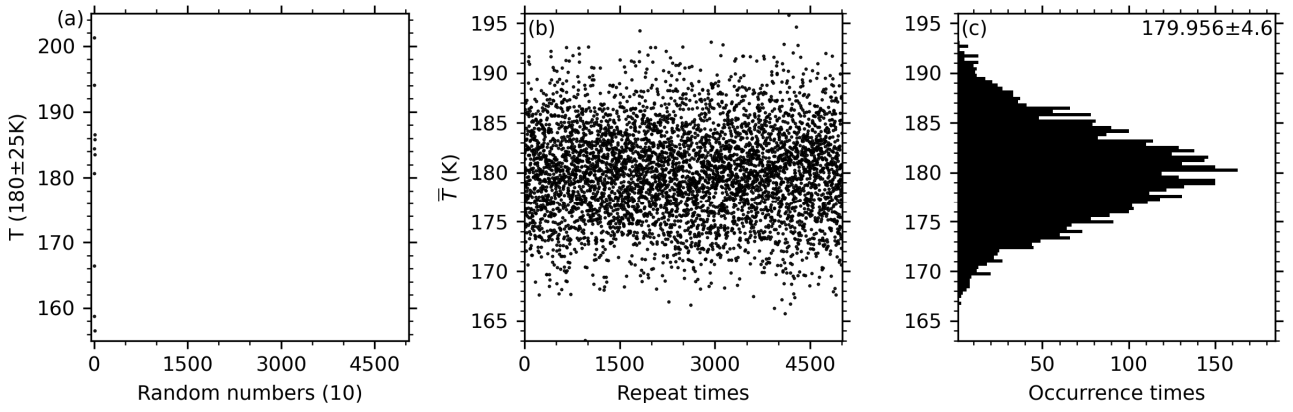
$$F_n(x) = P\left\{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \leq x\right\},$$
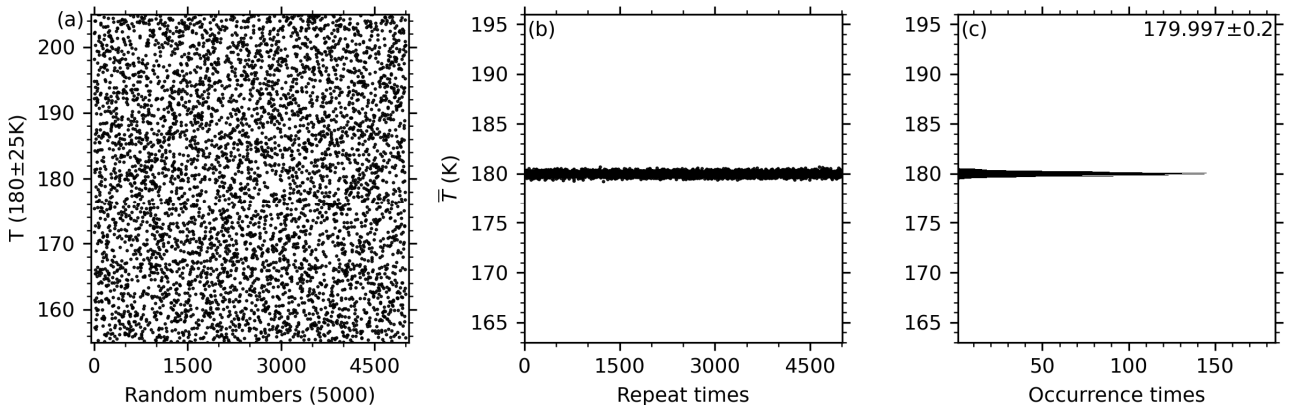
has limitation of,

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P\left\{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \leq x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

The CLT states that if one take sufficiently large samples from a population, the samples' means will be normally distributed, even if the population isn't normally distributed. Thus, X follows the normal distribution of $N(\mu, \sigma/\sqrt{n})$.

The uniform distribution in the range of $[a, b]$, its expectation and standard deviation are $\mu = (a + b)/2$ and $\sigma = (b - a)/\sqrt{12}$, respectively. According to CLT, the uncertainties of ±25K will induce an uncertainty of $50/\sqrt{12 \times 5000} \approx 50/245 = 0.204$K. This support the Monte Carlo simulations of ±0.20 K if there are 5000 samplings.

Figure R1. *Monte Carlo simulation on the influences of 10 samplings on the mean and standard deviation of the uniform distribution of ±25K with mean of 180 K. (a) an example of the 10 samplings; (b) and (c) show, respectively, the means and their histogram during 5000 times Monte Carlo simulations. The mean and standard deviation are labelled on the top right corner of (c).*

Figure R2. *Same caption as Fig. R1 but for 5000 samplings.*

3. The multiple linear regression equation contains terms involving the QBO. Have these been de-trended? Stratospheric temperature trends could create trends in the winds used in the QBO predictors. Failure to de-trend these predictors could lead to false or incorrect trends in the linear regression where the QBO predictors are significant.

Response: The QBO was not de-trended but retained its original form in our analysis. We agree that there might be trends in QBO and other predictors. To clarify this point, the followings were included in Sec. 2.3:

Here we note that both the trends (linear variations) and quasi-periodical variations represent the natural variations in QBO and other predictors. These natural variations might influence the trends and variations of temperature. Thus, MLR is applied to characterize the contributions from the natural variations of predictors, and then the resulted trends of temperature exclude the trends inhibited in the predictors. This is the trends studied in this work. Otherwise, if these predictors are de-trended,

their residuals are used in the MLR. The resulted temperature trends may include the trends inhibited in predictors.