Response to Reviewer # 1

(1) General comments:

The manuscript by Lu et al., submitted to ACP, presents observationally-based evidence of Sudden Stratospheric Warmings (SSWs) influence on air quality (AQ) over the Beijing-Tianjin-Hebei (BTH) region. The authors use reanalyses and observations fields over the 1980-2021 period to analyse stratosphere-troposphere coupling during 17 SSWs events, distinguishing between displacement (8) and split (9) type events. Responses of surface AQ over the BTH region, albeit substantially limited by the number of SSWs and (likely) large natural variability, show consistent changes with regard to current stratosphere-troposphere coupling understanding. This work extends and supports previous studies findings by including a number of SSW events, and finds stronger AQ responses over the BTH region associated with split SSW events.

The present study is of interest to the atmospheric chemistry and physics community. However, the language of the manuscript could be substantially improved, which would help to achieve a clear logic flow. The aim of the paper is clearly defined and technical concepts are usually presented. The methodology is sensible; however, it lacks a more detailed description to clearly understand the significance of the findings and their implications (see below). Results are well presented and support the conclusions. However, the analysis can greatly benefit from an additional discussion about the robustness of the results, e.g., limitations due to a relatively small number of SSW events and how this could be addressed in future work. I would recommend the present study for publication after the comments below are addressed.

Response: Thank you for your comments. We have included all of your suggestions in our revision, which help to improve the overall quality of our manuscripts.

(2) Specific comments:

The manuscript would greatly benefit from a substantial language revision (i.e., technical). The manuscript in its current form makes difficult the reading and understanding of the analysis. For example, please avoid using parenthesis to highlight the opposite (i.e., it takes at least two readings of the sentence to understand the meaning).

Response: The author has made a lot of technical revisions to the language of the manuscript.

We avoid using parenthesis to state the opposite position this time.

“According to the principle of geostrophic winds, decelerated westerly jets in the circumpolar region are accompanied with rise of polar cap height and/or the reduce of

the midlatitude heights, and accelerated winds are accompanied with decrease of polar height and/or rise of midlatitude heights." (L189-192)

The statistical test (i.e., t-test) used to define the significance of the results needs to be more clearly described. This involves the assumptions made, such as the distribution of the samples (e.g., normal or Student's t), and the variance of the samples (e.g., equal?), as well as the test used (i.e., one-sided or two-sided). All these details will greatly influence the statistical significance of the findings (e.g., Krzywinski & Altman, 2013; doi:10.0.4.14/nmeth.2698).

Response: The statistical test used to define the significance of the results has been clearly described this time.

"One-sample two-sided *t*-test is to test whether the difference between a sample average and a known population average is significant. When the population distribution is normal, but the samples are not large, the deviation of the sample mean from the population mean show a *t*-distribution. As the SSW samples are limited, it is reasonably to use the Student's *t*-test for this study. The one-sample *t*-test is calculated as $t = \frac{\bar{x}-\mu}{\frac{\sigma_x}{\sqrt{n}}}$, where n is the sample number, $\bar{x}$ is the sample mean, and σ_x is the sample standard deviation. The null hypothesis is that the *t*-value is zero (i.e., $\bar{x}=\mu$) if the sample mean shows insignificant difference from the population mean. Otherwise, the sample mean is significant different from the population mean if the null hypothesis is rejected (Krzywinski and Altman, 2013). In order to test the credibility and consistency of data, the bootstrap method is adopted to calculate the confidence level (e.g., Alfons et al., 2022) for the mean visibility, haze days, fog days, light fog days by resampling 1000 times with a sample size proportion of 0.5 for both displacement and split SSWs." (L137-146)

The authors often describe their results without addressing whether the signals are significant or not (i.e., yet significance is provided in their analysis). It is important to focus on relevant features (i.e., statistically significant) to understand the implications of the results.

Response: We considered your suggestion and mentioned the significance level in several places.

- "However, only the easterly anomalies above 50 hPa from day -10 to day 20 are statistically significant and reach the maximum (-30 m/s) around the onset dates (Fig. 1b)" (L161-162)
- "However, only the easterly anomalies above 200 hPa from day -15 to day 40 are statistically significant (Fig. 1e)." (L165-166)
- "No significant zonal-mean temperature anomalies are observed in the troposphere…" (L180)

- "In contrast, for the split SSWs, the positive geopotential height anomalies appear since day -22 and propagate downward to the troposphere instantly, where significant signals are observed." (L198-199)
- "…and the significant positive height center is located around the Bering Strait and North Pacific (Fig. 2a)." (L215-216)

Please see our revised manuscript for more details.

The manuscript lacks a discussion about the robustness and limitations of the results, e.g., taking into account that only 8-9 events for displacements and split types are accounted, respectively; what else can be done in future work to gain confidence?

Response: We added some discussion this time. Thank you very much!

"The composite results in this work are consistent with previous case studies, although only 8 or 9 SSWs are considered. In the future, outputs from climate-chemistry coupled models (e.g., Liang et al., 2022; Rao et al., 2022) can be used to further improve the robustness of diagnostic results with a much larger SSW sample size." (L502-505)

(3) Technical corrections:

L30. References only refer to AQ health impacts. Please, include/expand references that support haze pollution impacts on the ecological environment, transportation, and so forth.

Response: A reference about the impact of haze pollution on the ecological environment has been added. (L31)

L35-40. What is the Arctic Oscillation? Please describe (e.g., one sentence).

Response: "During the AO positive phase, the low pressure in the Arctic region deepens, and the high pressure in the midlatitudes intensifies, limiting the southward expansion of the cold air in the polar region.". (L41-43)

L63-64. Since this is a key feature of SSWs events, the authors may want to move the sentence at an earlier stage of the paragraph, i.e., "SSW event is a typical phenomenon of two-way coupling between stratosphere and troposphere (Hu et al., 2014)".

Response: The sentence has been moved to an earlier state of the paragraph. (L51-52)

L104-106. Observations provided by the China Meteorological Information Center need reference and/or link.

Response: Revised. (L109)

L137-141. Please clarify the meaning of the dates provided, e.g., onset of the SSWs events?

Response: Added "…and their onset dates are as follows". (L151, 153)

L148 and L152-154. Only for the split SSWs cases, right? With the analysis being not statistically significant below ~200hPa.

Response: Considering the comments from the other reviewer, the composite for all SSWs is removed.

- "When the displacement SSWs are considered, the zonal mean zonal wind anomalies at 60°N only appear above 200 hPa and do not propagate downward to lower troposphere. The easterly anomalies begin to appear 10 days before the onset dates and last until day 35. However, only the easterly anomalies above 50 hPa from day -10 to day 20 are statistically significant and reach the maximum (-30 m/s) around the onset dates (Fig. 1b)." (159-162)
- "In contrast, the easterly anomalies begin to appear since day -15 for split SSWs, and the wind anomaly magnitude is also stronger, reaching the maximum intensity (-35 m/s) several days before the SSW onset (Fig. 1e). The easterly anomaly signal can last until day 50, which propagate downward to the near surface since day 20. However, only the easterly anomalies above 200 hPa from day -15 to day 40 are statistically significant." (L162-166)

L166. Warmest anomalies for SSWs displacement events appear on day 5? Do you mean day -5?

Response: Changed. (L175)

L169-171. Such statement needs reference.

Response: Added a reference here. (L182)

L188-189. It is not clear what the sentence is referring to. Please clarify.

Response: Added "…for the split SSWs…" (L198)

L194-196. "... might have a downward influence on the tropospheric circulation. ..." This is a bit confusing given the statement in L199-200 ("... the stratosphere-troposphere coupling in the extratropics, which is followed by discernable tropospheric circulation anomalies...").

Response: Changed to "…which has a downward influence on the tropospheric circulation anomalies." (L204-205)

L200-201. I do not understand this sentence here. Are the authors not focusing on describing stratospheric circulation anomalies?

Response: Deleted “and troposphere”. (L211)

L233. PNA has not been introduced.

Response: Added. (L256-258)

L250. "The enhanced wavenumber 2 can also propagate upward into the stratosphere and split the polar vortex" needs a reference.

Response: Added. (L278)

L257. "are shown Fig. 4" -> "are shown in Fig. 4"

Response: Changed. (L283)

L284-287. The SSWs stratosphere-troposphere coupling needs reference.

Response: Added two references here. (L314)

Figure 7. The authors may want to include confidence intervals of the environmental metrics shown in the figure.

Response: Added. (L401-405)

L423. "... is more significant..." -> "... stronger..."?

Response: Changed. (L477)

L440. "Compared with the some previous..." -> "Compared to previous..."

Response: Changed. (L495)