*We thank both the reviewers for their comments on this manuscript. We have now revised the manuscript based on the reviewers' comments and provided a point-by-point response to their comments. The reviewers' comments are in blue, and our responses follow below and are prefaced by "Reply". Corresponding changes in the manuscript are also provided with section and line numbers. We also provide the track change version of the manuscript. The figures in the track change file are replaced by the revised figures and text changes are highlighted in blue color.*

## Reviewer 1

This study addresses changes in their haze weather index HWI. Much more evidence is needed that this HWI will correlate with future pollution/visibility. As mentioned briefly in the conclusions the actual changes in visibility will depend on other factors, especially emissions of particulate matter. Hence to avoid misinterpretation by readers "hazy days" and "clear days" need to be replaced throughout by high and low HWI.

**Reply:** We thank the reviewer for raising this point. Our study solely focuses on the meteorological changes leading to haze for a fixed high emission scenario and we have not examined the influence of emission reductions in the manuscript. We have now clearly stated this in the revised manuscript in the abstract and introduction as below. We find the usage of low and high HWI complicated as a reader. Therefore, we have reworded 'hazy days' to 'haze conducive weather' and 'clear days' to 'clear weather' and added HWI values in brackets in the revised text. Some examples are given in the lines below. We have preserved the usage of hazy and clear days in Section 3 as the terms are used in direct reference to air quality, i.e. $PM_{2.5}$ concentrations/visibility.

Abstract:
*Lines 20-24: We use a large-scale meteorology-based Haze Weather Index (HWI) with values >1 as a proxy for hazy conducive weather and HWI <-1 for clear weather conditions over the NCP. The PPE generated using the UK Met Office HadGEM-GC3 model shows that under a high-emission (RCP8.5) scenario, the frequency of haze conducive weather (HWI>1) is likely to increase whereas the frequency of clear weather (HWI<-1) is likely to decrease in future.*

*Introduction (section 1):*
*Lines 138-149: In this paper, our focus is on the daily clear or haze conducive weather conditions over the NCP under a fixed high-emission scenario (RCP8.5). For this purpose, we use the HWI proposed by Cai et al. (2018) as past research studies have shown a robust correlation between the HWI, which is a large-scale meteorology based index, and haze conducive weather for Beijing in China. Whilst Cai et al. (2018) originally proposed HWI for Beijing, the index is based on changes in large-scale meteorology over the NCP and thus offers a good potential as the indicator of haze conducive weather over the NCP. One potential advantage of using the HWI for future projections, as opposed to a regional or local air stagnation index, is that the general circulation models generally simulate large-scale meteorology reasonably well as compared to local or regional meteorology. Therefore, we expect the future projections of clear or haze conducive weather provided using the HWI to be less uncertain than projections provided using regional stagnation indexes.*

*Lines 159-168: In this paper, we first examine the application of the HWI as a proxy for haze conducive and clear weather over NCP for the current climate using a suite of observations (Section 3). We then provide the projections of the haze conducive (HWI >1) and clear weather (HWI <-1) frequency over NCP for the historical and future period.*

This study should include sufficient explanation for the HWI in this paper for the reader to understand the underlying principles without having to look into the earlier literature. It is not immediately obvious why these particular parameters should be associated with low haze, and why such specific regions are used to calculate the terms. Presumably, it is due to typical synoptic patterns bringing air from a clean or polluted origin (it would be helpful to provide illustrations of typical synoptic patterns associated with high and low HWI). This explanation is particularly necessary because the variable are much less directly associated with pollution than typical stagnation indices which usually based on boundary layer height, surface wind speed etc.

**Reply:** We thank the reviewer for the suggestions. We have now added a new figure on climatological wind speed in the lower and middle troposphere over the North China Plain (Fig.1 of the revised manuscript) and added text to explain the role of different variables considered in this study (introduction: lines 56-71, section 2.2: lines 220-229). We have also added a new figure and accompanying text on synoptic circulation patterns for hazy days with high $PM_{2.5}$ concentration and clear days with low $PM_{2.5}$ concentration (Fig. 2 of the revised manuscript). The impact of including more variables has already been examined by Cai et al. and is now also mentioned in the introduction (lines 150-155).

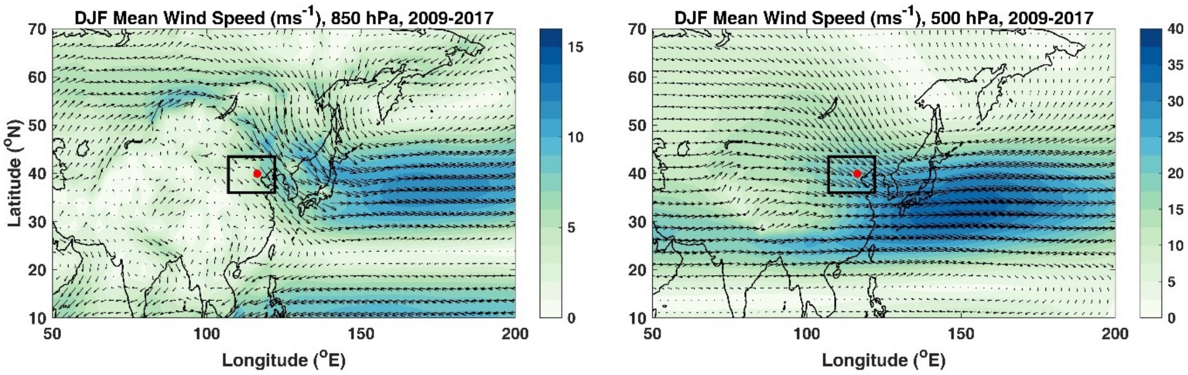

*__Figure 1__ Average wind speed at (a) 850 hPa and (b) 500 hPa pressure level. The red dot represents the location of Beijing and black rectangle shows the location of the NCP. This figure has been repeated for a longer average period, i.e. 1979-2019 (not shown) and the result is similar.*

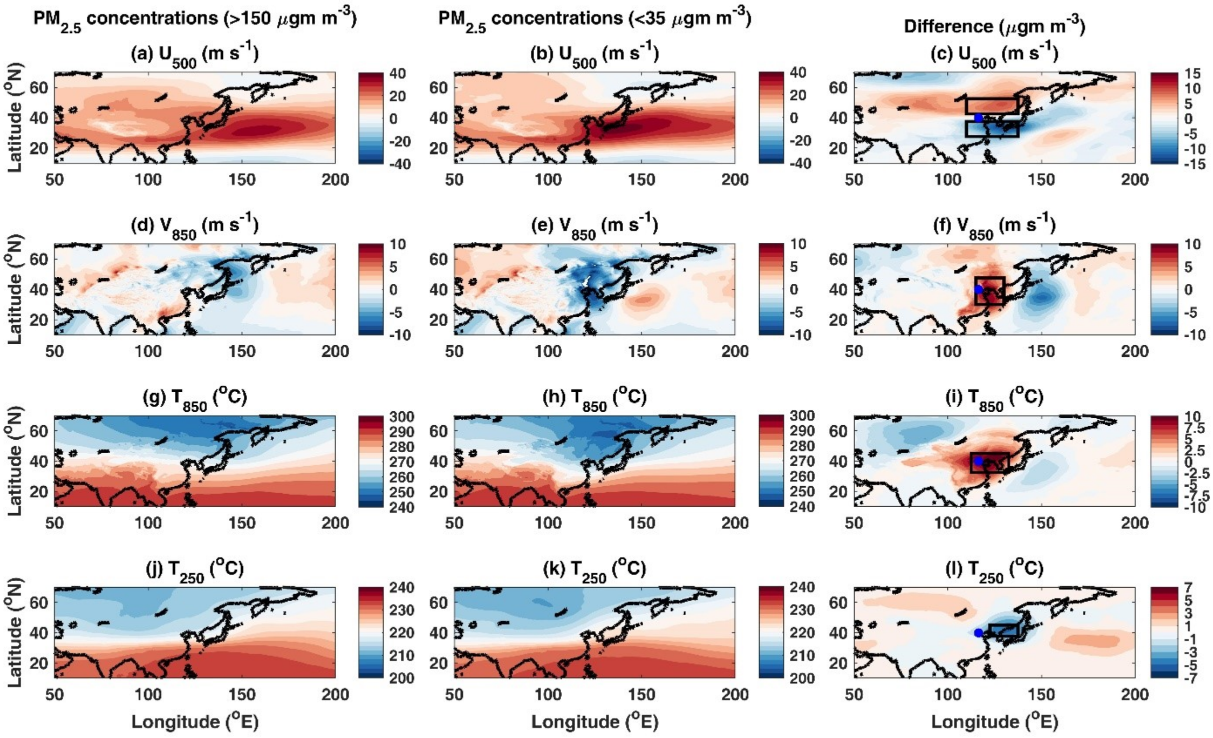

*Figure 2 Winter composites of u-wind at 500 hPa level ($U_{500}$) over China for all available days for which data is available from US embassy station for Beijing for DJF 2009-2017 for* ***(a)*** *high $PM_{2.5}$ (>150 μgm $m^{-3}$),* ***(b)*** *low $PM_{2.5}$ (<35 μgm $m^{-3}$) concentrations and* ***(c)*** *difference between the composites in (a) and (b).* ***(d-f)*** *same as (a-c) but for v-wind at 850 hPa level ($V_{850}$),* ***(g-i)*** *same as (a-c) but for temperature at 850 hPa level ($T_{850}$), and* ***(j-l)*** *same as (a-c) but for temperature at 250 hPa pressure level ($T_{250}$). Black rectangles (B1-B5) in the last column show the regions for which spatial means were used for the calculation of the HWI. The blue dot in these columns shows the location of Beijing.*

*Introduction:*

*Lines 56-71: The normal winter meteorological conditions over the NCP are characterized by northwesterly flow near the surface through to the mid-troposphere associated with the East Asian winter monsoon circulation (Fig. 1a and 1b; also see An et al., 2019; Renhe et al., 2014; Li et al., 2016; Xu et al., 2006; Chen and Wang, 2015). The northwesterly winds support the intrusion of relatively clean air from the high latitudes to the NCP and therefore ventilate this region (Xu et al., 2006). However, during the severe haze episodes, the near-surface northwesterlies appear to be weaker than normal and the mid-tropospheric trough was reported to be shallower and shifted northwards – collectively leading to a weaker than normal northwesterly flow and reduced horizontal transport of air pollutants from the NCP (Fig. 2a-b). In addition to changes in horizontal winds, the vertical temperature gradient between the lower and upper troposphere over the NCP can influence the vertical dispersion of the pollutants. A warmer than normal temperature near the surface, accompanied with colder temperature in the upper troposphere, would enhance the thermal stability and reduce the atmospheric mixing leading to the build-up of the atmospheric pollutants over this region (Fig. 2; also see Hou and Wu, 2016; Sun et al., 2014; Wang et al., 2014a; Zhang et al., 2018; Cai et al., 2018).*

*Lines 150-155: The HWI uses four meteorological variables as stated above, but Cai et al. (2018) have also examined the impact of the inclusion of more weather variables, such as* *geopotential height, boundary layer thickness and local stratification instability, in the HWI and did not find any significant differences in the performance of the HWI. Therefore, we use the same variables and methodology as Cai et al (2018) to calculate the HWI and provide future projections of haze conducive and clear weather using the HWI.*

Section 2.2:

Lines 220-229: *During the hazy days, the mid-tropospheric westerly flow becomes weaker over the NCP as compared to the clear days (Fig. 2a-c). The mid-tropospheric trough also moves northwards as suggested by the dipole pattern in Fig 2c, which shows the differences in the $U_{500}$ for hazy and clear days. The northerly flow near the surface is weaker during hazy days as compared to clear days (Fig. 2d-f). The lower troposphere is relatively warmer during hazy days as compared to clear days (Fig. 2g-i) whereas the upper troposphere is cooler over the NCP (Fig. 2j-l). The changes in these variables are also consistent with the previous studies (e.g. Cai et al., 2017) that showed similar changes for this time period. Therefore, we use these four variables for the calculation of the HWI, which is used as a proxy for haze conducive and clear weather conditions under a future climate.*

It is also not at all obvious that the same correlations between variables and pollution/visibility will hold in a future climate – as one example any systematic changes in humidity or boundary layer height over NCP with climate change would strongly affect haziness but not appear in the HWI. This may have been addressed in the studies cited, if so the specific examples need to be referred to, if not then the authors need to demonstrate themselves that HWI is applicable in future climates.

**Reply:** The reviewer has raised a good point. We would like to clarify that we use the correlation between HWI and air quality indicators only to show the influence of large-scale meteorology on air quality over Beijing and NCP. We use the HWI as a proxy for haze conducive weather conditions that lead to air pollution build-up under the current climate. We do not use HWI as a proxy for air pollution directly. These points are now clarified in the revised manuscript. Some examples of the revised text are given below. As the reviewer mentioned, our analysis is based on the underlying assumption that present meteorology that influences air pollution over the NCP in the current climate will also hold in future. We have mentioned this assumption in the introduction (lines 156-161).

Abstract:

Lines 19-21: *We use a large-scale meteorology-based Haze Weather Index (HWI) with values >1 as a proxy for haze conducive weather and HWI <-1 for clear weather conditions over the NCP*

Introduction:

Lines 138-139: *In this paper, our focus is on the daily haze conducive and clear weather conditions over the NCP under a fixed high-emission scenario (RCP8.5).*

Lines 153-158: *Therefore, we use the same variables and methodology as Cai et al (2018) to calculate the HWI and provide future projections of haze conducive and clear weather using the HWI. However, our analysis is based on an underlying assumption that the large-scale meteorological conditions, which are used as a basis for the HWI, will have a similar influence on the air quality of the NCP in the future climate as for present-day climate.*

*Lines 177-179: Daily mean $PM_{2.5}$ concentrations are constructed using hourly data to evaluate the performance of the HWI as a representative of haze conducive and clear weather conditions for Beijing (see Section 3).*

*Lines 255-259: As the HWI was originally proposed for Beijing by Cai et al. (2018), we first determine if the HWI can be used as a representative of haze conducive and clear weather conditions for the present climate for Beijing using (a) $PM_{2.5}$ concentrations from the US embassy station in Beijing and (b) $PM_{2.5}$ concentrations averaged over larger Beijing domain from CAQRA reanalysis and (c) visibility data from the CMA stations in Beijing.*

*Lines 344-345: Therefore, we use HWI > 1 as a proxy for haze conducive weather and HWI < -1 as a proxy for clear weather across the NCP region.*

A much stronger result could be obtained if any variables relating to pollution (aerosols or passive tracers) in the model were output. The authors should check what is available.

**Reply:** Our analysis is based on daily variables and the model outputs for passive tracers or aerosols is not available for daily time-scale. Also, please see our above response on the use HWI as a proxy for clear or haze conducive weather conditions and not of air pollution.

For the analysis in section 5 it is essential for the earlier section 3 to determine how much each of the variables contributes to the correlations with pollution/visibility. i.e. if V850 turns out to be the most important variable determining the correlation with pollution/visibility then the lack of future change in V850 would imply the haziness is unlikely to change in future.

**Reply:** The reviewer makes an interesting point. In this analysis, we show that the HWI has a robust influence on air quality over the NCP in the present climate. We have compared the correlation of each independent variable with the $PM_{2.5}$ concentrations observed at the US embassy station in Beijing for 2009-2017. The variables combined in the HWI index shows a correlation of ~0.6 with the $PM_{2.5}$ concentrations for this station. T250 shows a correlation of -0.4, T850 shows 0.59, U500 shows 0.34 and V850 shows 0.61. Similar values are also obtained for PM2.5 concentration data from CAQRA for 2013-2017. Statistically, the correlations for V850 and T850 are of similar magnitude whereas, the correlation between T250 and U500 is slightly weaker. Physically, multiple favourable weather conditions, as represented by each of these variables, collectively provide a conducive setting for haze. Hence, we focus on the HWI as a combined index rather than its individual components. This is now mention in Section 3.1 (Lines 285-290).

We now emphasise that the future projection of haze conducive and clear weather conditions are under the assumption that the HWI will have a similar influence on air quality in future climate as for present climate in the introduction (lines 155-158). We have merged the previous text in section 6 into section 4 and rephrased the sentences (lines 482-495).

*Introduction:*
*Lines 155-158: However, our analysis is based on an underlying assumption that the large-scale meteorological conditions, which are used as a basis for the HWI, will have a similar influence on the air quality of the NCP in the future climate as present climate.*

*Section 3.1:*
*Lines 285-290: We have also examined the relationship between the individual variables in the HWI (section 2.2) and $PM_{2.5}$ concentrations observed at the US embassy in Beijing/CAQRA and find that the individual components have correlation values that are similar to or less than*

*that of those used in the combined HWI. Also, physically multiple favourable weather conditions, as represented by each of these variables, collectively provide a conducive setting for haze. Hence, we focus on the HWI as a combined index rather than its individual components.*

*Section 4:*
*Lines 482-495:* *We now investigate changes in the distribution of the HWI as well as individual constituents of the HWI between the far future (2060-86) and the historical (1979-2005) period. The probability distribution of the HWI shows a shift in the distribution towards higher magnitudes for the far future as compared to the historical period (Fig. 8). This implies an increased frequency of haze conducive weather, as the number of days with HWI >1 increase. A similar shift is apparent in the zonal-mean wind ($U_{500}$) and the vertical temperature profiles (dT), whereas no apparent shift is noted in $V_{850}$. We also find that the shift in the HWI, as well as U500 and dT distribution, is not due to the shift in one particular PPE member or time period. It is consistent across the 16 PPE members and is continual over time from the historical to the far-future period. Therefore, for the PPE analysed here, the changes in the haze conducive weather (HWI>1) is largely associated with the changes in the $U_{500}$ and dT, and $V_{850}$ appear to have a less important role. Despite using a multimodel ensemble and a different time period than used here, a similar result with a relatively larger shift in the PDFs of $U_{500}$ and dT as compared to $V_{850}$ can also be noted in the Cai et al. (2017).*

Line 31: The sentence "The future frequencies of winter hazy and clear days in the PPE are largely driven by changes in zonal-mean mid-tropospheric winds and the vertical temperature gradient over the NCP" is very misleading. The study does not assess at all that future frequencies of hazy and clear days are driven by changes in zonal winds and vertical temperature gradients, rather it finds that those are the components of the HWI that change. There is no evidence presented of the effects on haze itself.

**Reply:** We thank the reviewer for pointing out our unclear text. We have re-phrased our text in the abstract and in section 4:

*Abstract:*
*Lines 27-29:* *"The future frequencies of haze conducive weather (HWI>1) during winter are associated with changes in zonal-mean mid-tropospheric winds and the vertical temperature gradient over the NCP."*

*Section 4:*
*Lines 491-493:* *Therefore, for the PPE analysed here, the changes in the haze conducive weather (HWI>1) is largely associated with the changes in the $U_{500}$ and dT, and $V_{850}$ appear to have a less important role.*

Lines 59-67: This paragraph describes important quantities (PBL height, stability, humidity) that are all key for haze formation but don't appear in the HWI. Much more evidence is needed that the particular variables in HWI are related to the quantities responsible for haze. References to literature on other stagnation indices (e.g. Horton et al. 2014, Vautard et al. 2018, Garrido-Perez 2021) needs to be discussed here.

**Reply:** We agree with the reviewer that there are several other indexes that could be used for this analysis. We have added a discussion to the introduction on the large-scale meteorology based indexes (lines 76-93) and regional or local meteorology-based stagnation indexes (lines

94-114). We have added our motivation to use HWI and its potential advantage in the lines 138-149 of the revised manuscript.

Introduction:
Lines 76-93: *On a daily scale, past studies have examined the changes in haze conducive weather conditions over China under climate change scenarios using large-scale meteorology-based indexes. For example, Cai et al. (2017) have used four key variables, i.e. meridional wind at 850 hPa ($V_{850}$), zonal wind at 500 hPa ($U_{500}$), temperatures at 850 hPa ($T_{850}$) and 250 hPa ($T_{250}$) pressure levels to calculate a meteorology-based daily Haze Weather Index (HWI). They have projected a ~50% increase in the frequency of winter haze conducive weather conditions, similar to the January 2013 event, over Beijing in the future (2050-2099) as compared to the historical (1950-1999) period under the RCP8.5 scenario using 15 CMIP5 models. Using the HWI, Liu et al. (2019) projected a 6-9% increase in the winter haze frequency under 1.5° and 2° global warming, respectively based on 20 CMIP5 models whereas Qiu et al. (2020) projected a relatively high increase of 21% and 18% in severe winter haze episodes under 1.5° and 2° global warming, respectively using an ensemble of climate simulations from the Community Earth System Model 1 (CESM1) (Kay et al., 2015). Callahan and Mankin (2020) also used specific humidity, $V_{850}$, $T_{850}$ and temperatures at 1000 hPa to examine the haze favourable meteorology for Beijing, and found a 10-15% increase in winter haze conducive weather in CMIP5 multimodel and CESM large ensemble under 3° warming. These authors have also emphasized a large influence of internal variability in addition to anthropogenic forcing on future haze conducive weather over Beijing.*

Lines 94-114: *In addition to the large-scale meteorology based indexes, several other stagnation indices based on regional or local meteorological variables have also been used to determine the influence of anthropogenic climate change on haze conducive weather for China as well as global regions. Using minimum monthly mean wind speeds averaged over northwestern Europe, Vautard et al. (2018) suggested a potential increase in the frequency of stagnant conditions conducive to air pollution over northwest Europe; however, their results were sensitive to models used for the analysis. Horton et al. (2014) have used thresholds for the daily mean near-surface (10-m) wind speeds, mid-tropospheric (500 hPa) temperatures and accumulated precipitation to calculate the Air Stagnation Index (ASI) under RCP8.5 scenario using 15 CMIP5 models. They found an increase in air stagnation occurrence events leading to poor air quality by up to ~40 days per year over a majority of the tropics and sub-tropics. Han et al. (2017) examined indicators of haze pollution potential (e.g. horizontal transport, wet-deposition, ventilation conditions) using three regional climate simulations and projected a higher probability of haze pollution risk over the Beijing-Tianjin-Hebei region under the RCP4.5 scenario. Garrido-Perez et al. (2021) took a different approach as compared to analysing probabilistic projections and used the ASI to generate stagnation storylines, i.e. plausible and physically consistent scenarios of stagnation changes based on the response of remote drivers under climate change forcing, for Europe and the United States (US).*

Lines 138-149: *In this paper, our focus is on the daily haze conducive and weather conditions over the NCP under a fixed high-emission scenario (RCP8.5). For this purpose, we use the HWI proposed by Cai et al. (2018) as past research studies have shown a robust correlation between the HWI, which is a large-scale meteorology based index, and haze conducive weather for Beijing in China. Whilst Cai et al. (2018) originally proposed the HWI for Beijing, the index is based on changes in large-scale meteorology over the NCP and thus offers a good potential as the indicator of haze conducive weather over the NCP. One potential advantage of using the HWI for future projections, as opposed to regional or local air stagnation index, is that the general circulation models generally simulate large-scale meteorology reasonably well as* *compared to local or regional meteorology. Therefore, we expect the future projections of clear or haze conducive weather provided using the HWI to be less uncertain than projections provided using regional stagnation indexes.*

Lines 68-91: More information needs to be given in this paragraph as to what measures of haze are used in these different studies. Do they all use HWI, or do some use a stagnation index, do some use modelled PM or visibility?

**Reply:** We have now provided details of the index and variables used by each study and have discussed their key findings in the revised manuscript. Please see lines 76-114 provided in our response to preceding comment.

Section 2.1: More information is needed here on these observations. It seems that using the US embassy site could lead to extremely localised sensitivities to wind directions/conditions depending on where the embassy is situated in relation to the sources of pollution or even orientation with respect to street canyons in the neighbourhood. Figure 1 would be much more credible if it used a more spatially-averaged measure of PM2.5.

**Reply:** We thank the reviewer for this suggestion. We have examined the relationship between $PM_{2.5}$ concentrations averaged over a larger domain and the HWI (fig. 3b and 3f). We do not find any substantial differences in the result and think that our result is robust across different data sources presented here. We have added this text to section 3.1 (previously section 2.1) at lines 291-302.

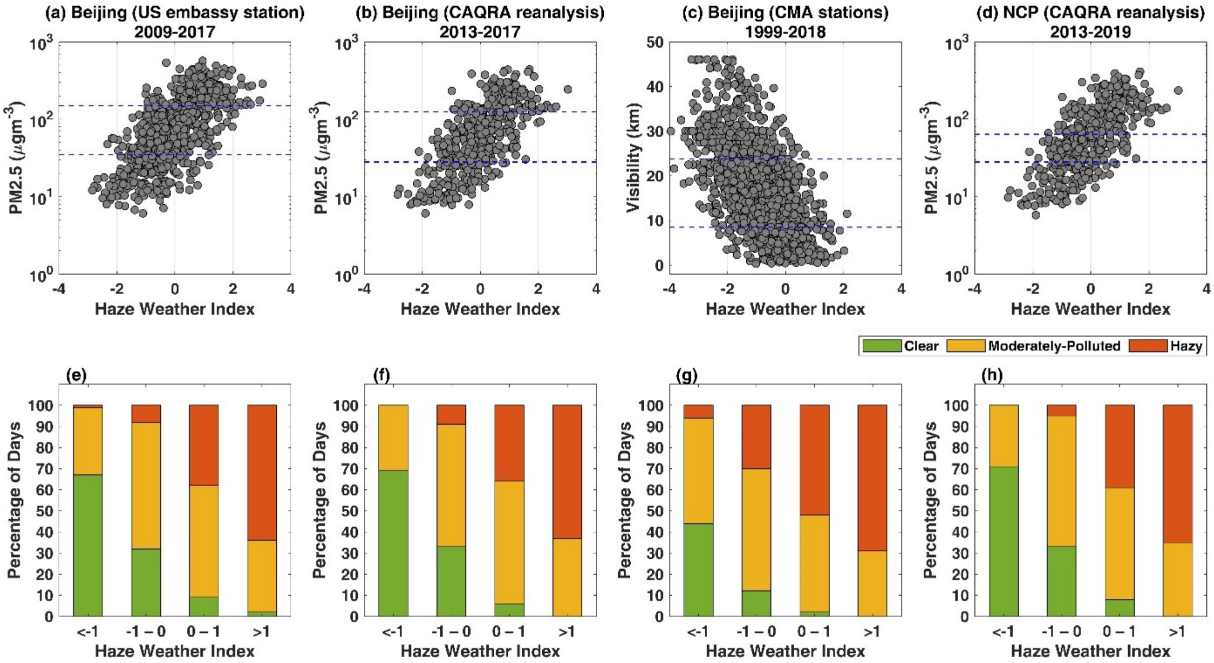

***Figure 3** HWI versus daily mean **(a)** $PM_{2.5}$ concentrations for the US embassy Beijing station for DJF 2009-2017 **(b)** $PM_{2.5}$ concentrations spatially averaged over the region around Beijing (116.15-116.65 °E, 39.65 - 40.15 °N) from CAQRA reanalysis for DJF 2013-2017 **(c)** visibility averaged over 20 stations from the CMA for DJF 1999-2018 and **(d)** $PM_{2.5}$ concentrations spatially averaged over the NCP (36-43.5 °N, 107-122 °E) from CAQRA reanalysis. Blue lines show the 25th and 75th percentile thresholds used to define clear and hazy days for each dataset. Percentage of clear, moderately polluted and hazy days for different HWI ranges for the **(e)** US embassy Beijing station for DJF 1999-2018 **(f)** larger Beijing domain (116.15-116.65 °E, 39.65 - 40.15 °N) from CAQRA reanalysis for DJF 2013-2017 **(g)** Beijing for DJF 1999-2018 **(h)** NCP from the CAQRA reanalysis for DJF 2013-2017.*

*Lines 291-302: To examine if the $PM_{2.5}$ concentrations from the US embassy station are sensitive to the abrupt changes in the local meteorology, e.g. wind speeds or direction, we also examine the relationship between the HWI and $PM_{2.5}$ concentrations averaged over the domain centred around Beijing (116.15 – 116.65 °E, 39.65 – 40.15 °N) from the CAQRA reanalysis data (Fig. 3b and 3f). The $PM_{2.5}$ concentrations for region spatially averaged around Beijing from CAQRA data are in the range 6 $\mu g\ m^{-3}$ – 441 $\mu g\ m^{-3}$ and from the Beijing US embassy station are 6 $\mu g\ m^{-3}$ – 569 $\mu g\ m^{-3}$ suggesting the values from both data sources are comparable. The correlation coefficient is ~0.58, which is the same as the correlation obtained using the US embassy data. The total number of hazy, clear and moderately polluted days for different HWI ranges also show similar results for both datasets (Fig. 3e-3f). This implies that the HWI relationship with $PM_{2.5}$ concentrations is robust across different data sources and that $PM_{2.5}$ is a regional pollutant.*

Lines 252-258: This seems a strange comparison, there is no reason why the average PM2.5 in two different regions over two different time periods should be the same. It would be much more useful to show a correlation between Beijing and NCP to see whether Beijing is typical.
**Reply:** We thank the reviewer for pointing this out. We have added another figure and accompanying text to section 3.1 which correlates the PM2.5 concentrations for larger Beijing domain from the CAQRA reanalysis with the HWI (please see our response to the preceding comment).

Figure 5: It would be useful here to look at the absolute frequencies as well as the change wrt historical. For instance there seems to be an outlier with very high changes starting in 2006-2032, but this might be explained if the historical frequency happened to be particularly low. These points all need error bars to determine whether the differences between members are explained by internal variability. I couldn't see any of the pick/red circles that "suggests a decrease in daily haze frequency".
**Reply:** The reviewer has made excellent suggestions. We have now replaced this figure with new Figure 6a (see below), which shows the absolute frequencies for HWI>1 and HWI<-1 in the revised manuscript. We have unintentionally used wrong averaging period for near future and that is why there was an outlier. This has now been corrected in the revised manuscript text– we thank the reviewer for pointing this out.

We have now shown the impact of anthropogenic climate change and parametric effect on the future projections of mean frequencies in section 4 (lines 414-456). We have now also added more details on the ensemble member showing reduction in daily haze conducive weather frequency at lines 385-392.

*Section 4:*
*Lines 385-392: It is noted that, for all three periods, only one of the sixteen ensemble members (E16 shown in Fig. 10) shows a reduction in the haze conducive weather frequency whereas other ensemble members show an increase in frequency for all periods. For the historical period, the E16 ensemble member has a mean frequency of 16.3, which reduces to 16.2, 14.4 and 15.2 for near, mid and far future. While E16 ensemble member shows a consistent reduction in mean frequency in future, the reduction is specific to only this ensemble member and is not a general feature across PPE members.*

*Lines 414-456: We also determine the influence of anthropogenic climate change and the parametric effect on the frequencies of haze conducive weather (HWI>1) and clear weather (HWI<-1) for the historical as well as the three future periods. As shown in later Section 5, the* *estimate of interannual variance from the control is representative of all time periods and shows no discernible parametric effect. Therefore, we pool the 16 PPE control simulations to sample the internal variability for box and whiskers shown in Fig. 6 (a) and 6 (b) (see captions for details on resampling).*

*In Fig. 6 (a), we show the mean frequency of haze conducive weather and clear weather for 16 individual PPE members (circles) and PPE mean (triangles). The grey box and whiskers represent the range of ensemble mean frequencies that can be explained by the internal variability. If the PPE mean (triangles) lies within the whiskers (i.e. 95 percentile of the control distribution) we conclude no influence of anthropogenic climate change on mean frequency however if the PPE mean lies outside the whiskers, it would represent a climate change signal in the mean frequency. Figure 6 (a) suggest that the mean frequencies for haze conducive as well as clear weather lies within the box-whiskers for the historical but lies outside the whiskers for the three future periods, thereby showing a clear impact of anthropogenic climate change on the frequencies of both haze conducive and clear weather conditions.*

*We now examine whether the differences in the mean frequency across different PPE members (shown by circles in Fig. 6a) for a given period can be explained by the internal variability or if the differences in PPE members partly arise due to the parametric effect. The triangles in Fig. 6b shows the variance across 16 PPE members, i.e. variance across 16 circles shown in Fig. 6a, for each time period. The whiskers in Fig. 6b show the 95th confidence interval from the control simulation and is representative of the internal variability. For any time period, if the PPE member variance (triangle) lies within the whiskers, we conclude that the differences in mean frequencies in Fig. 6a can be fully explained by the internal variability and there is no discernible impact of the parametric effect. However, if the triangles lie outside the whiskers in Fig. 6b, we conclude an impact of the parametric effect on the mean frequency for that period. For the points that lie outside the whiskers in Fig. 6b, we also quantify the percentage of variance that can be explained by the internal variability and parametric effect. For any time period, the variance in ensemble mean due to the parametric effect is simply calculated as follow and the remaining variance is attributed to the internal variability.*

$$\frac{\text{Total variance in the ensemble mean} - \text{Mean variance from the control simulation}}{\text{Total variance in the ensemble mean}} \times 100$$

*Figure 6b shows that the variance in PPE mean frequency for historical and future periods lies within the range sampled by the internal variability for both haze conducive weather (HWI>1) and clear weather (HWI<-1). For mid-future, the variance in haze conducive weather lies outside the whiskers and whereas the variance for clear weather lies within the whiskers. For mid-future and for haze conducive weather, the internal variability can explain ~33% of the variance across PPE members and the remaining ~67% arises due to the parametric effect. For the far future, triangles corresponding to both haze conducive and clear weather lies well outside the whiskers and therefore show a clear influence of parametric effect. Only ~20% of the variance in the frequency of haze conducive weather and ~43% variance in the frequency of clear weather can be explained by the internal variability and the remaining 80% and 57% respective variance in the frequencies arise due to the parametric effect.*

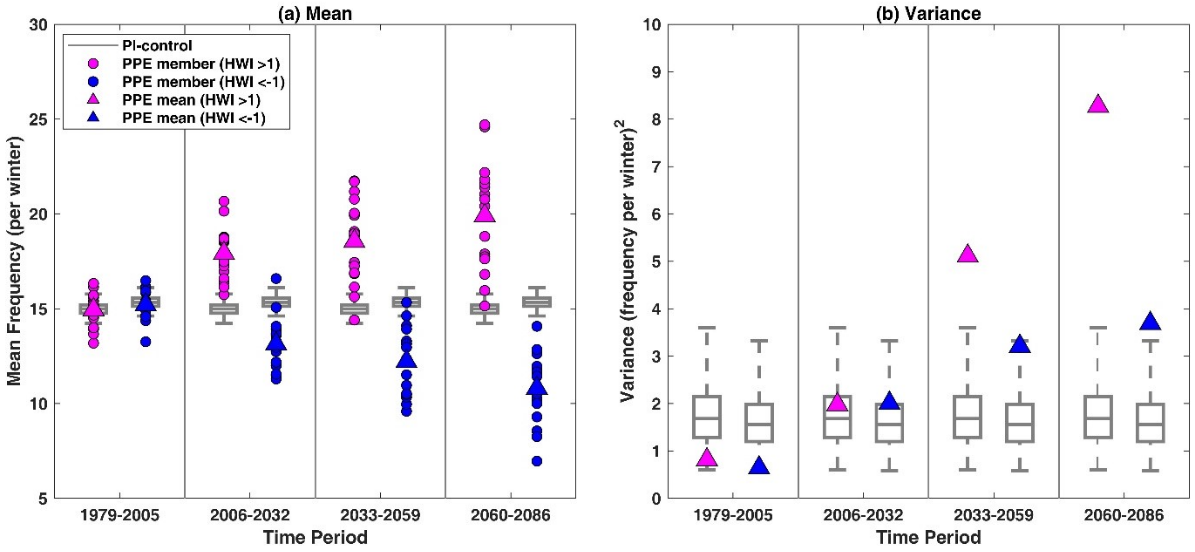

*Figure 6 (a) Mean frequency of haze conducive weather (HWI>1, pink) and clear weather (HWI<-1, blue) for the historical period (1979-2005), near (2006-2032), mid (2033-2059) and far (2060-2086) future under the RCP8.5 scenario. Circles represent PPE members and triangles PPE mean. Grey box and whiskers show the distribution of 10,000 values of mean frequencies sub-sampled from the control simulation, (b) same as (a) but shows variance across 16 PPE members for each period. For box and whiskers, we first randomly sampled 10,000 time series of length 27 years using 2704 years of pre-industrial control simulation and calculated 10,000 values of mean frequency. We then randomly sub-sample 16 mean values (corresponding to the number of ensemble members) from the 10,000 mean values, calculated their mean for (a) and variance for (b). This is repeated 10,000 to obtain a distribution. The boxes are at the 25th and 75th percentile and the whiskers at 2.5th and 97.5th percentile of mean and variance distribution. For panel (a), the box and whiskers are comparable only to the ensemble means (triangles) and not ensemble members (circles).*

Line 347-349: I didn't quite follow how the shift in HWI,U500 and dT were "consistent" when different PPE members show large difference in HWI in figure 5.

**Reply:** We meant consistent across members and years, which means the PDF is not shifted because of one anomalous time period or ensemble member but all members show a gradual shift over the time period. We have now clarified this in section 4 (lines 487-490).

<u>Lines 488-490:</u> *We also find that the shift in the HWI, as well as U500 and dT distribution, is not due to the shift in one particular PPE member or time period. It is consistent across the 16 PPE members and is continual over time from the historical to the far-future period.*

Lines 437-438: This sentence isn't clear. The different PPE members certainly show different trends, but is the argument that these differences in trends can be explained by internal variability? The outlier member in figure 5 seems to suggest that for this particular set of parameters the trend is significantly higher than the others.

**Reply:** We use the control simulations to disentangle the impact of anthropogenic climate change and of the parametric effect from the internal variability. If the variance in trends across different PPE members lie within the 95th percentile distribution of the trends resampled using the control simulation, we conclude that the internal variability can explain the differences across the members. The time periods for which the PPE variance emerges outside the 95th percentile distribution of the internal variability, we have also added the percentage of the variance frequencies that can be explained by the internal variability. These points are now clarified in lines 414-456 provided in our response to above comments.

Figure 10. Why does the 2060-2086 circle in panel (c) show a negative trend in frequency when figure 5 shows all the trends are positive?
**Reply:** We now provide an explanation on this in following lines:

*Section 6:*
*Lines 594-599:* *Figure 11 (a) also shows a positive mean trend in haze conducive weather (HWI>1) for historical, near and mid future, but a weak negative trend for far future. While the frequency of haze conducive weather increases for all three future periods with respect to the historical period as shown in Fig. 6a, the trends only show an increment or reduction for that period as these are not referenced to the historical period. Therefore, trends could still be negative within any selected period, as in the case of the far future.*

Lines 465-466. This sentence of the climate change impact only being discernible for specific periods needs to be explained more fully. If it is true that an increase in HWI with climate is not a general feature then that disagrees with the claims of a " clear impact of anthropogenic climate change on future trends".
**Reply:** We have now rephrased this sentence as follows:

Section 7:
*Lines 658-666:* *We find a consistently growing influence of anthropogenic climate change and parametric effect on the mean haze conducive and clear weather frequencies across the 21st century. This suggests that in addition to the internal variability, the parametric effect adds as an additional source of uncertainty in future projections of haze conducive and clear weather, particularly towards the end of the 21st century. We find that the impact of anthropogenic climate change is discernible in trends for the historical period for haze conducive weather and up to mid of the 21st century for clear weather. Beyond these periods, the historical trends are not sustained and not distinguishable from the internal variability.*

## Reviewer 2

1. Line 133. "in (Kong et al., 2021)". Rephrase it.

**Reply:** This has been corrected in the revised manuscript:

*Lines 185-186:* *More details on the validation of the CAQRA dataset against the independent station data is provided in Kong et al. (2021).*

2. Line 135. "Chinese Meteorological Agency". Check and confirm this organization.

**Reply:** The organization name has been corrected in the revised manuscript – apologies for our oversight.

*Lines 186-188:* *The visibility data for Beijing (homogenized data for 20 stations in Beijing) is provided by the National Meteorological Information Center of China, China Meteorological Administration (CMA), for DJF 1999-2018.*

3. Lines 155-156. "(see next section for the cut-offs values used for PM2.5 concentration)". The thresholds for haze and clear weather days are given in lines 153-154.

**Reply:** We have now rephrased this sentence as follows

<u>Line 208:</u> *"see section 3.1 for the explanation on the $PM_{2.5}$ concentration cut-offs values used here"*

4. Lines 156-158. These statements could be made more clearly. Suggest adding the climatological zonal and meridional winds in panel a and b to illustrate the relationships between wind anomalies and the climatology.

**Reply:** We agree with the reviewer and we have now added climatological mean winter circulations in Fig. 1 of the revised manuscript. We have also added composite means of each variable for high HWI and low HWI in Fig. 2 of the revised manuscript. We have also added text corresponding to this figure in Section 1 (lines 56-71) and Section 2.2 (lines 220-229).

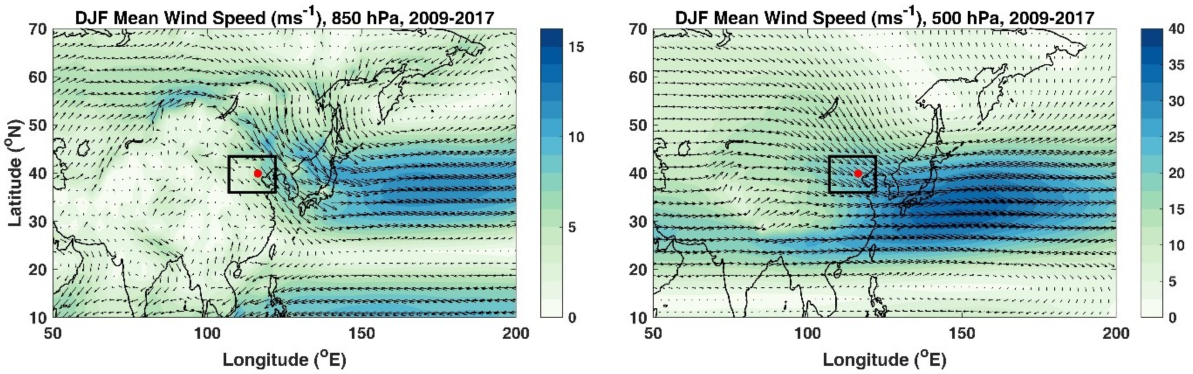

***Figure 1*** *Average wind speed at (a) 850 hPa and (b) 500 hPa pressure level. The red dot represents the location of Beijing and black rectangle shows the location of the NCP. This figure has been repeated for a longer average period, i.e. 1979-2019 (not shown) and the result is similar.*

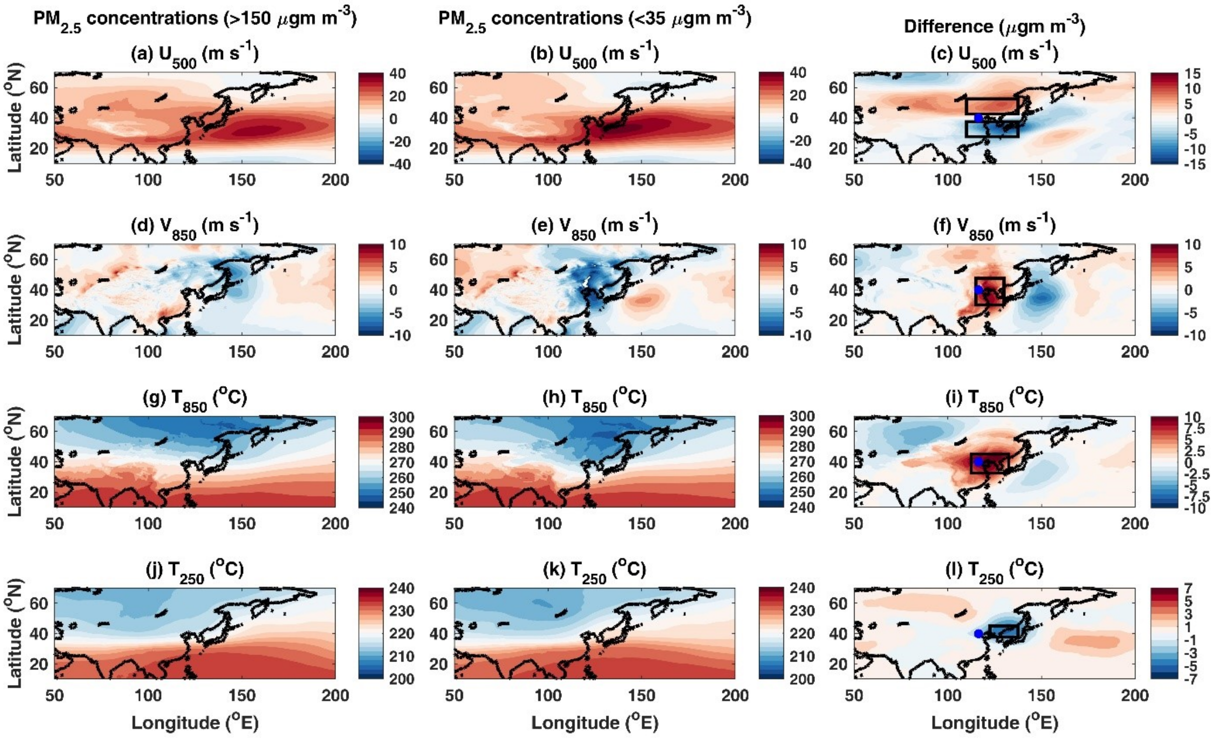

***Figure 2*** *Winter composites of u-wind at 500 hPa level ($U_{500}$) over China for all available days for which data is available from US embassy station for Beijing for DJF 2009-2017 for* ***(a)*** *high $PM_{2.5}$ (>150 µgm $m^{-3}$),* ***(b)*** *low $PM_{2.5}$ (<35 µgm $m^{-3}$) concentrations and* ***(c)*** *difference between the composites in (a) and (b).* ***(d-f)*** *same as (a-c) but for v-wind at 850 hPa level ($V_{850}$),* ***(g-i)*** *same as (a-c) but for temperature at 850 hPa level ($T_{850}$), and* ***(j-l)*** *same as (a-c) but for temperature at 250 hPa pressure level ($T_{250}$). Black rectangles (B1-B5) in the last column show the regions for which spatial means were used for the calculation of the HWI. The blue dot in these columns shows the location of Beijing.*

*Introduction:*
*Lines 56-71: The normal winter meteorological conditions over the NCP are characterized by northwesterly flow near the surface through to the mid-troposphere associated with the East Asian winter monsoon circulation (Fig. 1a and 1b; also see An et al., 2019; Renhe et al., 2014; Li et al., 2016; Xu et al., 2006; Chen and Wang, 2015). The northwesterly winds support the intrusion of relatively clean air from the high latitudes to the NCP and therefore ventilate this region (Xu et al., 2006). However, during the severe haze episodes, the near-surface northwesterlies appear to be weaker than normal and the mid-tropospheric trough was reported to be shallower and shifted northwards – collectively leading to a weaker than normal northwesterly flow and reduced horizontal transport of air pollutants from the NCP (Fig. 2a-b). In addition to changes in horizontal winds, the vertical temperature gradient between the lower and upper troposphere over the NCP can influence the vertical dispersion of the pollutants. A warmer than normal temperature near the surface, accompanied with colder temperature in the upper troposphere, would enhance the thermal stability and reduce the atmospheric mixing leading to the build-up of the atmospheric pollutants over this region (Fig. 2; also see Hou and Wu, 2016; Sun et al., 2014; Wang et al., 2014a; Zhang et al., 2018; Cai et al., 2018).*

Section 2.2:
*Lines 220-229: During the hazy days, the mid-tropospheric westerly flow becomes weaker over the NCP as compared to the clear days (Fig. 2a-c). The mid-tropospheric trough also moves northwards as suggested by the dipole pattern in Fig 2c, which shows the differences in the $U_{500}$ for hazy and clear days. The northerly flow near the surface is weaker during hazy days as compared to clear days (Fig. 2d-f). The lower troposphere is relatively warmer during hazy days as compared to clear days (Fig. 2g-i) whereas the upper troposphere is cooler over the NCP (Fig. 2j-l). The changes in these variables are also consistent with the previous studies (e.g. Cai et al., 2017) that showed similar changes for this time period. Therefore, we use these four variables for the calculation of the HWI, which is used as a proxy for haze conducive and clear weather conditions under a future climate.*

5. Lines 181-182. It would be helpful for readers if authors can give a brief description of the meaning the HWI.

**Reply:** We agree with the reviewer and we have now added more information on the HWI in the revised manuscript. Some examples follow:

*Lines 220-229: During the high haze days, mid tropospheric westerly flow becomes weaker over the NCP as compared to the low haze days (Fig. 2a-c). The mid tropospheric trough also move northwards as suggested by the dipole pattern in Fig 2c, which shows the differences in the $U_{500}$ for hazy and clear days. The northerly flow near the surface is weaker during hazy days as compared to the clear days (Fig. 2d-f). The lower troposphere is relatively warmer during hazy days as compared to the clear days (Fig. 2g-i) whereas upper troposphere is cooler over the NCP (Fig. 2j-l). The changes in these variables is also consistent with the previous* *studies (e.g. Cai et al., 2017) that showed similar changes for this time period. Therefore, we use these four variables for the calculation of the HWI, which is used as a proxy for clear and haze conducive weather conditions in future climate.*

*Lines 342-345:* *Overall, our results confirm that the daily HWI has a robust relationship with daily $PM_{2.5}$ concentrations not only for the Beijing station but across the NCP for the given time periods. Therefore, we use HWI > 1 as a proxy for haze conducive weather and HWI < -1 as a proxy of clear weather across the NCP region.*

6. Labels in Figure 2 are too small to read.

**Reply:** We have revised this figure (now Figure 3) and increased the size of labels.

7. Lines 268 and 269. Insert "weather" between "clear" and "conditions".

**Reply:** We have now rephrased this sentence as follows in the revised manuscript.
*Lines 347-347:* *"We now calculate the frequency of haze conducive weather (>1) and clear weather (HWI <-1) for the past and future using ERA-5 reanalysis and PPE members"*

8. Lines 311-314. Very long sentence.

**Reply:** We have now split this sentence into two sentences in the revised manuscript.
*Lines 410-413:* *"Overall, most ensemble members show an increase in the frequency of haze conducive weather and a reduction in the frequency of clear weather for all three future periods. However, negligible change or even the opposite change, though less likely, but possible for all periods."*

9. Line 253. Figure 7. It is not very clear what is the x-axis for those PDFs.

**Reply:** We have now added details on x-axis in caption of this figure (now Figure 8).

*Lines 499-502:* *The PDF for the HWI is created using the daily DJF time series of all 16 PPE members. PDFs for $V_{850}$, $U_{500}$ and dT are created using the normalized daily DJF time series of each variable calculated for the HWI (see section 2.2 for details) and represents the constituent variables of the HWI.*