****

In this file, **the text in black** shows the comments from reviewers and editor, while **the text in blue** is our replies.

****

SUMMARY:

Using observation data and reanalysis data, this paper utilizes a quantitative measurement of cold airmass to identify CAO and related dynamic/thermodynamic properties. The authors find the generic existence of air pollution reappearance after CAO, and raise a possible mechanism in the manuscript. This manuscript overall is interesting, but I do have some comments regarding the details of data and methods. My major and minor concerns are described below.

Response: Thank you very much for your kind evaluation. All of your comments are accepted and revised accordingly. We consider that your comments have helped greatly to improve the manuscript. The detailed response and revision are given below.

MAJOR COMMENTS:

1. The authors define north China as 30-40 N, 114-122 E. However, this is not a good definition. Furthermore, two sounding stations used in this manuscript, Nanjing and Baoshan (which is in Shanghai), are in East China. The authors can refer to the definition of the North China Plain in Kang et al. 2018

Kang, S., Eltahir, E.A.B. North China Plain threatened by deadly heatwaves due to climate change and irrigation. Nat Commun. 9, 2894,2018. https://doi.org/10.1038/s41467-018-05252-y

Response: We agree that the spatial range of North China Plain is usually defined as 113°–121°E, 34°–41°N (Kang and Eltahir, 2018), which is slightly north than the definition in our study. Here, we would like to explain why we select the region 114°–122°E, 30°–40°N as the study area. The determination of study area is based on the characteristics of air pollution and cold air activity. Figure R1 shows that the air pollution (AQI > 100) mainly occurs between the latitudes of 30°N and 40°N. The cold airmass activities also can affect areas around 30°N. In addition, our previous study shows the air quality at 30°N (Nanjing and Shanghai) usually varies following (only a half day's delay) the changes of air quality at 40°N during a CAO (Figure 7 in Liu et al., 2019). This suggest the changes of air quality in area between 30°N and 40°N could be considered as a whole.

In addition, Figure 1 in Kang and Eltahir (2018) shows the spatial distribution of topography, area equipped for irrigation, highest daily maximum wet-bulb temperature and population density. The characteristics of these variables, as well as AQI and cold airmass depth in Figure R1, in areas 30°–34°N and 34°–40°N are highly consistent, indicating the area 30°–40°N (Black boxes) is more appropriate to be selected as the study region.

Following your suggestion, we refer to the study region (114°–122°E, 30°–40°N) as "Northern and Eastern China" (NEC). We also add some explanations of the study region in Section 2 at Lines 85–88: "The determination of study area (NEC), including parts of both northern China and eastern China, is based on the characteristics of both air pollution and cold air activity. The air pollution (AQI > 100) mainly occurs between the latitudes of 30°N and 40°N, and the CAOs also usually affect areas around 30°N (figure omitted)."

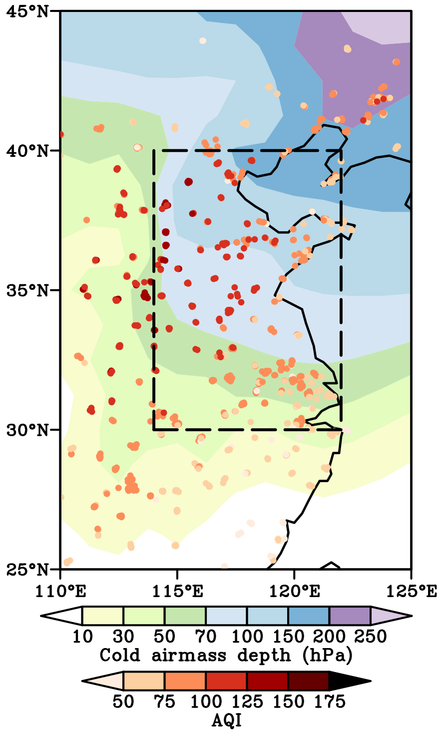

Figure R1. Winter mean cold airmass depth (shading) and AQI (colored dots) during 2014/2015–2021/2022.

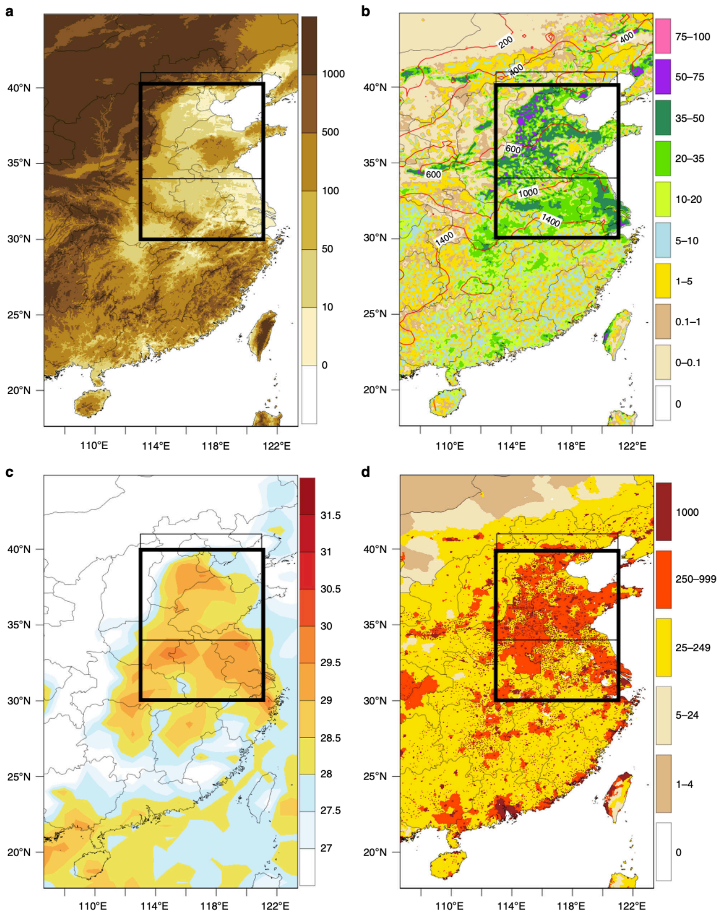

Figure 1 in Kang and Eltahir 2018. Brief characterization of Eastern China. Spatial distribution of a topographic map (m), b area equipped for irrigation (AEI, %) for 2005 from Historical Irrigation Data with climatology of annual precipitation from TRMM (contour, mm) in modern record (1998–2015), c highest daily maximum wet-bulb temperature from ERA-Interim, TW_{max} (°C) in modern record (1979–2016), and d population density in people/km^2. The box in each plot indicates the North China Plain used for regional analysis in this study.

References

Liu Q, Chen G and Iwasaki T 2019 Quantifying the impacts of cold airmass on aerosol concentrations over North China using isentropic analysis *J. Geophys. Res. Atmos.* **124** 7308–7326

2. In Section 2.1, more details are needed. For example

(1) What's the spatial distribution of those local AQI stations? Are most of them in the big cities?

Response: Figure R2 shows the spatial distribution of AQI stations used in our study. The stations have a relative even distribution in the study area. Some of the AQI stations are located in big cities (shaded by dark red) and some of them are in rural/suburban region (shaded by light yellow). According to your comments, we add description about the AQI observations at Lines 88–89: "The AQI stations have a relative even distribution in the study area (figure omitted). The stations are located in not only urban areas but also suburban and rural areas."

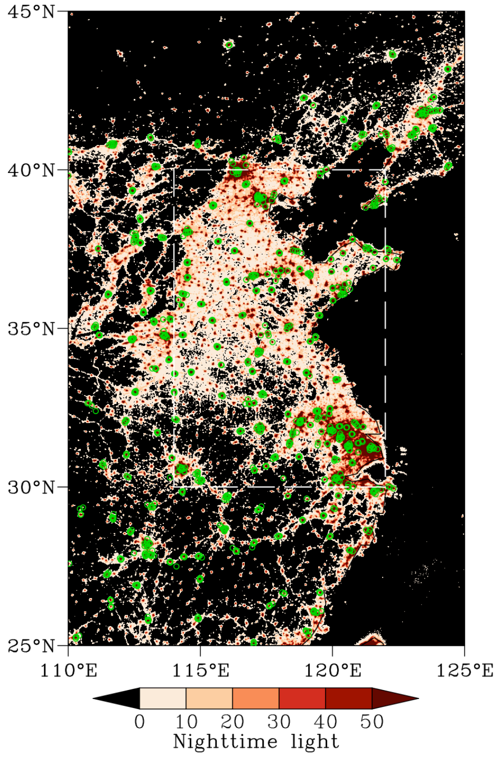

Figure R2. Spatial distribution of AQI stations (green circles) and nighttime light (shading) in 2013. The nighttime light values range from 0-63 and large (small) value denotes the urban (rural/suburban). The white box denotes the northern and eastern China. (Data source https://www.ngdc.noaa.gov/eog/dmsp.html).

(2) What's the vertical and horizontal resolution for the sounding data? And why do authors use sounding station data for the wind and air temperature? The temporal resolution is not good. Why don't you use reanalysis data as well?

Response: The sounding data used in our study has a higher vertical resolution than reanalysis data especially in atmospheric boundary layer. For example, sounding data at Beijing station has 60–70 levels in total and 8–14 levels below 850 hPa during study period of 14–17 Dec 2016 (Figure 7a in manuscript), while the reanalysis data has 37 vertical levels with 7 levels below 850 hPa (see reply to MAJOR COMMENT #2(3)). For spatial resolution, there are not many available radiosonde stations in the study area. So, we only selected four stations distributed from north to south. We add some of above descriptions of sounding data at Lines 95–97: "The sounding data used in our study usually has a higher vertical resolution than mainstream reanalysis data especially in atmospheric boundary layer. For example, sounding data at Beijing station has 60–70 levels in total and 8–14 levels below 850 hPa during a CAO event of 14–17 Dec 2016."

The sounding data provides a direct detection of atmospheric vertical profiles, which is more reliable than the model-simulated reanalysis data. Some studies also show that the boundary layer height calculated by reanalysis data has some biases as compared to sounding observation (Guo *et al*., 2021). Moreover, the atmospheric boundary layer conditions calculated by sounding observation data could well explain the changes of AQI observations. Therefore, we need first to verify the reliability of reanalysis data with sounding data.

We agree with the reviewer's suggestion that reanalysis data may be a good substitute for sounding data, since reanalysis data has a higher spatial and temporal resolution. Indeed, we have used reanalysis data to analyze the boundary layer characteristics. In section 4.1, we found the depth of residual cold airmass calculated by reanalysis data is highly consistent with the mixing layer height calculated by sounding data. The depth of residual cold airmass and mixing layer height also has a physical connection as discussed in the manuscript at Lines 233–243. This result suggests the use of reanalysis data is appropriate and reliable in our study. In addition, the calculation of cold airmass depth using reanalysis data is much easier than the calculation of MLH using sounding. Therefore, we used the reanalysis data instead of sounding data in the rest part of section 4.

References

Guo J, Zhang J, Yang K, Liao H, Zhang S, Huang K, Lv Y, Shao J, Yu T, Tong B, Li J, Su T, Yim S H L, Stoffelen A, Zhai P and Xu X 2021 Investigation of near-global daytime boundary layer height using high-resolution radiosondes: first results and comparison with ERA5, MERRA-2, JRA-55, and NCEP-2 reanalyses *Atmos. Chem. Phys.* **21** 17079–17097

(3) I am not sure which JRA-55 products are used in this study. First of all, JRA-55 should have a 3-hourly reanalysis, and usually, the vertical pressure levels are 60 levels.

Response: In this study, we use the standard product of isobaric analysis data (6-hourly, 37 layers) of JRA-55 following previous studies on cold air outbreak (Kanno *et al*., 2015; Abdillah *et al*., 2017; Liu *et al*., 2021). This data is freely available at JMA (http://search.diasjp.net/en/dataset/JRA55) and NCAR (https://rda.ucar.edu/datasets/ds628.0/).

The JRA-55 product users' handbook can be found in this website (https://jra.kishou.go.jp/JRA-55/document/JRA-55_handbook_LL125_en.pdf). At page 8 of the handbook, the time interval of isobaric analysis data is described as: "These fields are produced every six hours at 00, 06, 12 and 18 UTC". Page 14 shows the vertical coordinates of the data in isobaric fields as follow: "Isobaric fields are produced for 37 isobaric surfaces

(1000, 975, 950, 925, 900, 875, 850, 825, 800, 775, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, 225, 200, 175, 150, 125, 100, 70, 50, 30, 20, 10, 7, 5, 3, 2 and 1 hPa) except dew-point depression (or deficit), specific humidity, relative humidity, cloud cover, cloud water, cloud liquid water and cloud ice, which are produced for 27 levels from 1000 to 100 hPa only".

Following your suggestion, we add some detailed descriptions at Lines 77–79: "To identify the CAO events, we used isobaric analysis data of Japanese 55-year reanalysis (JRA-55). This dataset is freely available at JMA (http://search.diasjp.net/en/dataset/JRA55) and NCAR (https://rda.ucar.edu/datasets/ds628.0/). The JRA-55 has a horizontal resolution of 1.25° with 37 vertical pressure levels and a time interval of 6 hours (00, 06, 12 and 18 UTC)."

References

Kanno Y, Abdillah M R and Iwasaki T 2015 Charge and discharge of polar cold air mass in northern hemispheric winter *Geophys. Res. Lett.* **42** 7187–7193

Abdillah M R, Kanno Y and Iwasaki T 2017 Tropical–extratropical interactions associated with East Asian cold air outbreaks. Part I: Interannual variability *J. Clim.* **30** 2989–3007

Liu Q, Chen G, Wang L, Kanno Y, and Iwasaki T 2021 Southward cold airmass flux of the East Asian winter monsoon: Diversity and impacts *J. Clim.* **34** 3239–3254

MINOR COMMENTS:

1. In the introduction, the recent COVID-19 lockdowns also provide a unique opportunity to study the complex chemical effects of air pollution as well as meteorology. Here are some references.

1) Le T, Wang Y, Liu L, Yang J, Yung YL, Li G, Seinfeld JH. Unexpected air pollution with marked emission reductions during the COVID-19 outbreak in China. Science. 2020 Aug 7;369(6504):702-6.

2) Wang Y, Wen Y, Wang Y, Zhang S, Zhang KM, Zheng H, Xing J, Wu Y, Hao J. Four-month changes in air quality during and after the COVID-19 lockdown in six megacities in China. Environmental Science & Technology Letters. 2020 Sep 9;7(11):802-8.

3) Zhao N, Wang G, Li G, Lang J, Zhang H. Air pollution episodes during the COVID-19 outbreak in the Beijing–Tianjin–Hebei region of China: an insight into the transport pathways and source distribution. Environmental Pollution. 2020 Dec 1;267:115617.

Response: Following your suggestion, we add above references and some discussions in the introduction at Lines 36–37: "Even when local emission is obviously reduced during COVID-19 lockdown, severe air pollution still occurs in North China (Zhao et al., 2020)." and Lines 40–42: "The recent COVID-19 lockdowns also provide a unique opportunity for studying the complex chemical effects of air pollution as well as meteorology (Le et al., 2020; Wang et al., 2020)."

2. Section 2.1, "with observations made 24 times per day", is it measured equal frequency (i.e., 1h frequency)?

Response: Yes, the observation has a time interval of 1 hour. The revised text can be found at Line 84: "with observations of 1hour frequency".

3. Line 87, 89, vertical integral -> vertical integration.

Response: Revised as your suggestion.

4. For the definition of mass flux, the common definition is the rate of mass flow (SI unit kg/(m^2 s).

Response: We agree the unit of mass flux is usually expressed as kg/(m^2 s). In our study, we use the horizontal flux of cold airmass to describe the horizontal movement of cold air, which is different from the mass flux of cold air.

Following Iwasaki et al. (2014), the total cold airmass amount/depth (*DP*) is defined as the depth of air layer between ground surface (p_s) the isentropic surface of threshold potential temperature (θ_T), $DP = p_s - p(\theta_T)$ (area below the thick solid contour of 280K in Figure 2 in Iwasaki *et al* 2014). Thus, the definitions of dynamical and thermal properties are based on such definition of cold airmass (*DP*, unit: hPa). θ_T

Here, the isentropic mass continuity equation can be written as

$$\frac{\partial}{\partial t}\left(\frac{\partial p}{\partial \theta}\right) + \nabla \cdot \left(\frac{\partial p}{\partial \theta} v\right) + \frac{\partial}{\partial \theta}\left(\frac{\partial p}{\partial \theta}\dot{\theta}\right) = 0.$$

The vertical integration of above equation from surface boundary (θ_s) to θ_T leads us to total cold airmass conservation

$$\frac{\partial}{\partial t} DP = -\nabla \cdot \int_{p(\theta_T)}^{p_s} v dp + G(\theta_T),$$

where $\int_{p(\theta_T)}^{p_s} v dp$ is defined as the horizontal flux of cold airmass (*F*). Thus, the cold airmass flux (*F*) has a unit of hPa m s^{-1}.

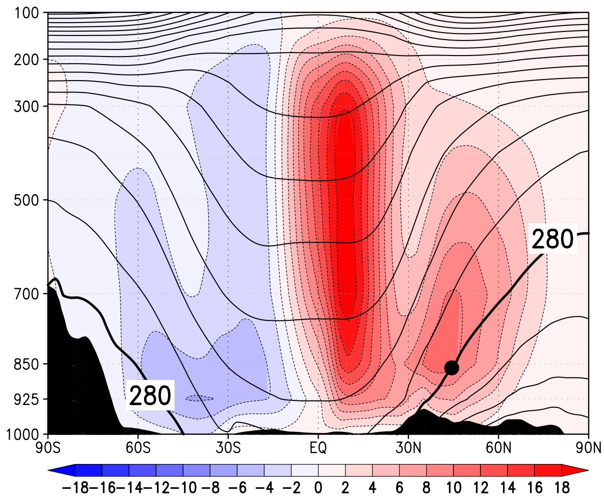

Figure 2 in Iwasaki *et al* 2014. Meridional cross sections of MIM's mass stream functions (contour interval 10^{10} kg s^{-1}) and potential temperature (contour interval 10 K) for DJF means over 1980/81–2009/10. A black dot is placed at the turning point of 45°N and isentropic zonal mean pressure of 850 hPa. Black colors indicate the zonally averaged topography.

5. Line 95, why do authors use the standard deviation to define the CAO? This only makes sense when the mean value of cold airmass depth is very close to zero. (Think about one example, if the cold airmass depth is 50 hPa, the standard deviation is still 169.7 hPa, then CAO should be defined as a cold airmass depth exceeding 50+169.7 = 219.7 hPa

Response: Sorry for the misleading description in the original manuscript. We fully agree that the selection of threshold value should consider the sum of standard deviation and mean value, which has been adopted in our study. The standard deviation and mean value of regional mean cold airmass are 78.6 hPa and 91.1 hPa, respectively. Thus, the threshold value of CAO is set as 78.6 + 91.1 = 169.7 hPa.

It should be noted that the study period has expanded to 2014/2015~2021/2022 in the revised manuscript. Thus, the threshold value of CAO becomes 166.8 hPa (i.e. 77.3+89.5 hPa). See revised text at Lines 112–114: "Thus, the CAO in this study is identified when the regional mean cold airmass depth exceeds 166.8 hPa, which is the sum of mean value (77.3 hPa) and standard deviation (89.5 hPa) of cold airmass depth on all winter days. According to the above criteria, 52 CAOs are identified over the 8 winters."

6. In Figure 1, The day 0 (March 09) are both in before-CAO and During-CAO periods. But Mar 10 is only in the During-CAO period. So, the definition of the period boundaries is not consistent. This may affect the analysis results (e.g., Figure 4 and so on).

Response: The day 0 is defined as the onset of a CAO event, which is the first day when regional mean cold airmass depth exceeds the threshold (166.8 hPa). Thus, for example, day +1 (−1) denotes 1 day after (before) the onset of the CAO event.

The definitions of different periods (before, during and after CAO) are intended to measure the reappearance of air pollution (see detailed definition in revised manuscript at Lines 112–118). We also correct the description of period during the CAO, which indeed starts from the onset day (day 0) and ends at the day when regional mean cold airmass depth falling below the threshold 166.8 hPa (Figure R3). To describe maximum and minimum values in the three periods more clearly and intuitively, we add some tags (AQI_b, AQI_d and AQI_a) in Figure R3. Results associated with the three periods are only shown in Figure 3 in the manuscript and will not affect the analysis results shown in Figure 4 and so on.

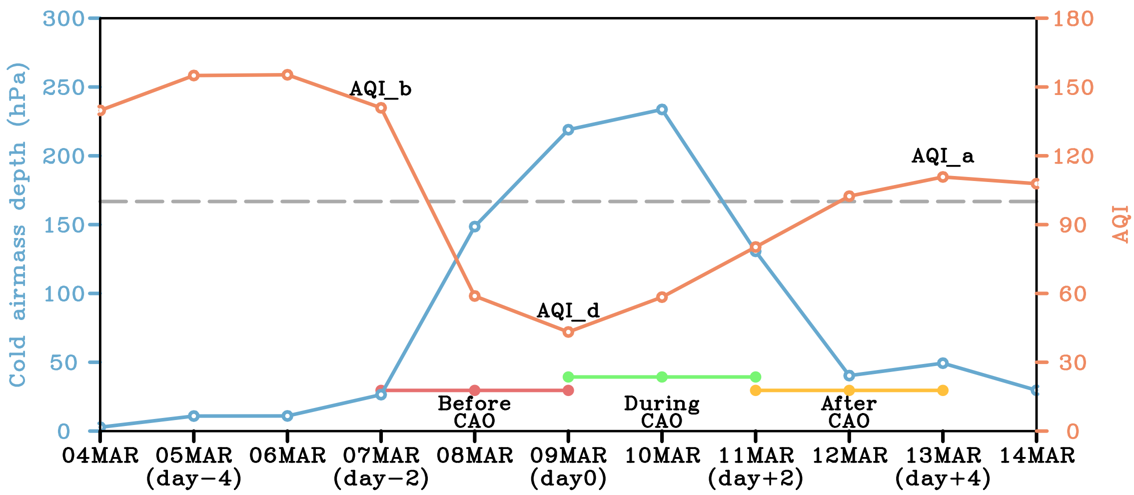

Figure R3. Time series of the regional averaged cold airmass depth (blue line) and AQI (orange line) in northern and eastern China (114°–122°E, 30°–40°N). The gray dashed line denotes the threshold value of the cold airmass depth.

In Figure 1/Figure R3, the maximum AQI in periods before (AQI_b) and after (AQI_a) CAO represent the original air pollution and reappeared air pollution, respectively. The minimum AQI during CAO (AQI_d) represent a relatively clean state. Ideally, the period before CAO should not include day 0, which is partly coincide with the period during CAO. However, in some CAO events, the maximum AQI before CAO-related-reduction (AQI_b) occurs in day 0 (Figure R4). This is because the cold airmass at day 0 may only influence the northern part of northern and eastern China, although the regional mean cold airmass depth exceeds the threshold value. Our previous study has also shown that air pollutants will increase before the cold airmass arrives, and reach the maximum when the cold airmass just begins to invade (Liu et al., 2019). Therefore, to accurately measure the air quality before the reduction caused by CAO, we extend the period before CAO to day 0.

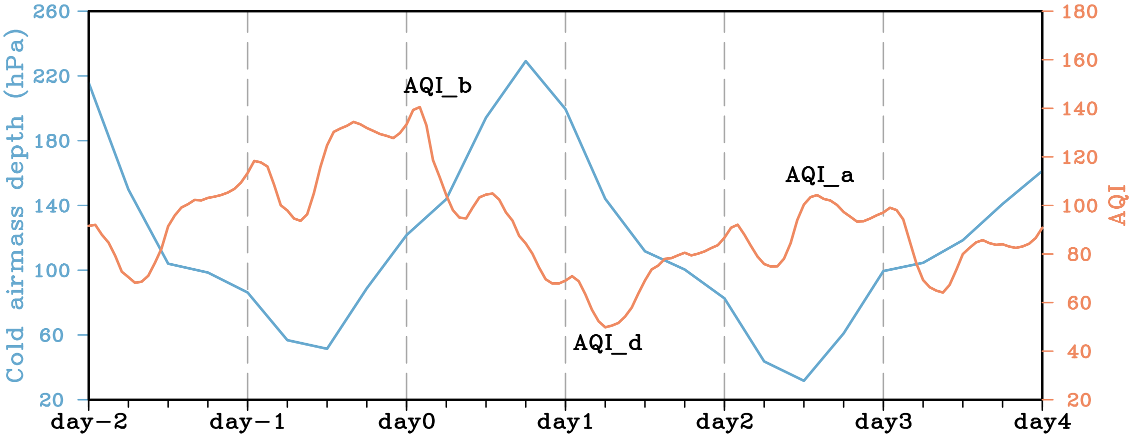

Figure R4. Evolutions of spatially averaged (northern and eastern China: 114°–122°E, 30°–40°N) cold airmass depth and AQI during a CAO event from 5 to 10 Dec 2017.

See revised text at Lines 116–121: "The onset of a CAO event, which is the first day when regional mean cold airmass depth exceeds the threshold (166.8 hPa), is described as the day 0. The period during the CAO starts from the onset day (day 0) and ends at the day when cold airmass depth falling below the threshold (day +2 in CAO event plotted in Figure 1). The period before CAO is defined as the two days before onset day to the onset day (days −2 to 0). The period after CAO, which is also called the decay phase, is defined as the three days after cold airmass depth falling below the threshold (days +2 to +4 in CAO event plotted in Figure 1)."

7. Figure 4, better to define the different color lines in the caption.

Response: Revised as your suggestion. See caption of Figure 4: "Blue lines and red lines denote the cold airmass depth and AQI, respectively".

We acknowledge your great help to improve the manuscript.

Thank you very much.