This manuscript addresses a timely and relevant topic about how meteorological conditions modulate the response of co-occurring PM2.5 and O3 pollution to precursor emission reductions, including the role of online aerosol-meteorology coupling. The response matrices and the attempt to distinguish five local meteorological regimes are potentially useful. However, the current manuscript overstates the robustness and generality of several conclusions. The framing of DHP versus separate pollutant responses needs to be sharpened, the effective sample size and uncertainty are not addressed, and the paired ARI/ACI experiment does not isolate the individual feedback pathways. In addition, the model has substantial PBLH, wind, and O3 biases that directly affect the central analysis. I therefore recommend major revision. Comments to the authors

Major comments

1. At line 37, the abstract motivates the study through NO_x*-VOC emission reductions, but the manuscript is framed as a study of double-high pollution (DHP), which requires PM2.5 and O3 to be high simultaneously. Reducing* NO_x *and VOCs primarily targets ozone chemistry;* NO_x *also affects nitrate aerosol, but SO2, BC, OC, primary particulate matter, NH3, and other particle relevant emissions are not reduced in the stated scenarios. Thus, the current experiments do not represent general "emission reductions" or a complete DHP control strategy. Please clearly reframe the work as the response of DHP to* NO_x*-VOC precursor perturbations. The abstract and conclusions should explain why a* NO_x*-VOC-only matrix is scientifically sufficient for the question being asked, or narrow the question accordingly.*

Reply: We thank the reviewer for this insightful comment.

We agree that the experiments perturb only NO_x and VOC emissions and do not represent a comprehensive control strategy that would also involve SO_2, primary particulate matter, NH_3, and other species. We have revised the manuscript to reframe the work accordingly and to justify why a NO_x–VOC-only perturbation matrix is scientifically sufficient for the specific question being asked. The reasons are as follows:

(1) O_3 formation is entirely governed by NO_x–VOC chemistry. Surface O_3 is produced through photochemical reactions involving NO_x and VOCs while other emitted species (SO_2, BC, OC, primary particulate matter, NH_3) do not directly participate in O_3 photochemistry. Therefore, any investigation of O_3 response to precursor controls must focus on NO_x and VOC.

(2) NO_x reductions directly target nitrate, the dominant $PM_{2.5}$ component during

DHP. Our model results show that for DHP samples, NO_3^- is the dominant $PM_{2.5}$ component, accounting for 34.0% of $PM_{2.5}$, far exceeding SO_4^{2-} (9.7%), NH_4^+ (13.6%), BC (4.4%), and OC (8.8%). This is consistent with Fig. S7, which already notes that "NO_3^- is the dominant secondary inorganic component". Because NO_x is the direct precursor of HNO_3 and thus NO_3^-, NO_x reductions effectively target the largest $PM_{2.5}$ component.

(3) NH_3 is not perturbed because it serves as the neutralizing agent for both NO_3^- and SO_4^{2-}. Under typical ammonia-rich conditions of summer BTHS, NO_3^- formation is HNO_3-limited rather than NH_3-limited (Li et al., 2018; Liu et al., 2021; Li et al., 2025), so reducing NO_x directly reduces the limiting precursor. In addition, SO_2, BC, OC, and primary PM are excluded from the perturbation design because their contributions to $PM_{2.5}$ are relatively small and they are photochemically inert, thus not influencing O_3. Including their reductions would confound the isolation of the NO_x–VOC sensitivity signal that is the central focus of this study.

(4) This design follows established practice in co-pollution control research. Previous studies examining synergistic $PM_{2.5}/O_3$ control have similarly focused on NO_x–VOC perturbation matrices (Ding et al., 2022; Qian and Liao, 2025; Liu et al., 2021), recognizing that NO_x and VOC are the common precursors linking the two pollutants.

Based on the above reply, we have replaced "emission reductions" with "NO_x–VOC precursor perturbations" in the title, Abstract, the research objective at the end of the Introduction, the scenario design description in Section 2.3, the title of Section 3.4 and 3.5, and the first paragraph of the Conclusion, while "emission reductions/controls" is retained where control effectiveness and management implications are discussed. The specific revisions are as follows:

(1) The title has been revised to "Meteorological conditions drive divergent responses of co-occurring $PM_{2.5}$ and O_3 pollution to NO_x–VOC precursor perturbations and the impact of aerosol-meteorology feedback in Beijing–Tianjin–Hebei–Shandong in July 2020".

(2) In the Abstract, the sentence "This study investigated the DHP response to NO_x–VOC emission reductions and the role of aerosol-meteorology feedback in Beijing–Tianjin–Hebei-Shandong" has been revised to "This study investigated the $PM_{2.5}/O_3$ response to NO_x–VOC precursor perturbations, which jointly control O_3 and nitrate, and the impact of aerosol-meteorology feedback on $PM_{2.5}/O_3$ response in

Beijing–Tianjin–Hebei–Shandong in July 2020 using WRF-Chem.”.

(3) The fifth paragraph of introduction has been revised to “However, how NO_x–VOC precursor perturbations affect DHP under refined local meteorological conditions remains unclear. Most previous studies focused on large-scale synoptic classifications or individual meteorological variables, which do not directly link to the chemical environment governing precursor reactions. In addition, the role of aerosol-meteorology feedback in affecting control effectiveness under different meteorological conditions is poorly understood, as most previous coupled studies considered ARI only and did not evaluate how this feedback affects $PM_{2.5}$/O_3 with the variations of boundary layer, thermal and moisture conditions. To fill these gaps, this study investigates DHP responses to NO_x–VOC precursor perturbations in Beijing–Tianjin–Hebei–Shandong (BTHS), a hotspot of summertime DHP, in July 2020. The typical meteorological conditions of DHP were identified and classified into different types based on combined boundary layer, temperature and humidity conditions rather than large scale synoptic categories or a single meteorological element. Then, two-dimensional response matrices were constructed using the online-coupled WRF-Chem model to evaluate O_3 and $PM_{2.5}$ responses to NO_x–VOC precursor perturbations and aerosol-meteorology feedback contribution under each meteorological condition type. The response was also linked to key meteorological and chemical indicators to elucidate the underlying mechanisms.”

(4) In the second paragraph of Section 2.3, we have added “These scenarios are designed to isolate the response of $PM_{2.5}$/O_3 to NO_x–VOC precursor perturbations. NO_x and VOC are the common precursors linking O_3 photochemistry and nitrate aerosol formation. Under the typical ammonia-rich conditions of summer BTHS, nitrate formation is HNO_3-limited (Li et al., 2018; Liu et al., 2021; Li et al., 2025), so NO_x reductions directly target the most responsive $PM_{2.5}$ component. Other emitted species (SO_2, BC, OC, primary particulate matter, NH_3) are held at baseline levels to avoid confounding the NO_x–VOC sensitivity signal.”.

(5) In the third paragraph of Section 3.3, we have added “For DHP samples, NO_3^- accounts for 34.0% of $PM_{2.5}$, exceeding SO_4^{2-} (9.7%), NH_4^+ (13.6%), BC (4.4%), and OC (8.8%), confirming that NO_x-driven NO_3^- is the dominant $PM_{2.5}$ component and the primary pathway through which NO_x reductions affect $PM_{2.5}$.”

(6) The titles of Section 3.4 and 3.5 have been revised to “3.4 Responses of DHP to precursor perturbations under different meteorological conditions” and “3.5 The

impact of aerosol-meteorology feedback on pollutant response to precursor perturbations".

(7) The first paragraph of Conclusion has been revised to "This study investigated the response of co-occurring $PM_{2.5}$ and O_3 pollution to NO_x–VOC precursor perturbations under typical meteorological conditions and the impact of aerosol-meteorology feedback on this response using WRF-Chem with 36 NO_x–VOC precursor perturbation scenarios (0–50% each) and paired simulations with and without aerosol-meteorology feedback for July 2020 over BTHS.".

(8) In the sixth paragraph of Conclusion, we have also added "Furthermore, the perturbation matrix isolates NO_x–VOC sensitivity and does not represent a comprehensive emission control strategy. Reductions of SO_2, NH_3, primary particulate matter, and other species would provide additional $PM_{2.5}$ co-benefits that are not quantified here.".

We have added the following references:

Li, H., Zhang, Q., Zheng, B., Chen, C., Wu, N., Guo, H., Zhang, Y., Zheng, Y., Li, X., He, K.: Nitrate-driven urban haze pollution during summertime over the North China Plain, Atmos. Chem. Phys., 18, 5293–5306, https://doi.org/10.5194/acp-18-5293-2018, 2018.

Liu, Z., Zhou, M., Chen, Y., Chen, D., Pan, Y., Song, T., Ji, D., Chen, Q., Zhang, L.: The nonlinear response of fine particulate matter pollution to ammonia emission reductions in North China, Environ. Res. Lett., 16, 034014, https://doi.org/10.1088/1748-9326/abdf86, 2021.

Li, L., Guo, Y., Xu, J., Ye, X., Li, D., Liu, Z., Ti, C., Liu, X., Zhang, L.: Opportunities to mitigate $PM_{2.5}$ and nitrogen deposition through agricultural NH_3 control strategies in China's Beijing–Tianjin–Hebei region, Environ. Sci. Technol., 59, 14572–14584, https://doi.org/10.1021/acs.est.5c02485, 2025.

2. At line 46, "aerosol feedback" is ambiguous. It could mean aerosol-radiation interaction (ARI), aerosol-cloud interaction (ACI), the resulting changes in temperature, PBLH, humidity, and clouds, or the subsequent effect of those meteorological changes on DHP. Please define the term in the abstract and introduction using a clear causal chain. The paper should distinguish direct aerosol effects on radiation/cloud microphysics from the meteorological response and from the final change in PM2.5/O3 or joint DHP occurrence. It would also help to state whether the

intended quantity is the feedback contribution to pollutant concentrations, the feedback contribution to the emission-control response, or the feedback contribution to the probability of DHP. At present, the abstract and Sect. 3.5 mainly discuss PM2.5 and O3 separately, so the reader cannot tell how the "aerosol feedback" changes the DHP.

Reply: We thank the reviewer for this insightful comment.

In the revised manuscript, "aerosol feedback" is replaced throughout by "aerosol-meteorology feedback" and defined as a causal chain in the Abstract and Introduction. We have also stated that the impact of aerosol-meteorology feedback on the response of $PM_{2.5}/O_3$ concentration to precursor perturbations is the intended quantity, denoted $\Delta_{aer}PM_{2.5}/O_3$ in Section 2.3.

We have added Fig. S15 to show the impact of aerosol-meteorology feedback ($\Delta_{aer}X$) on the response of aerosol optical depth at 550 nm (AOD_{550}), shortwave downward radiation (SWDOWN, W m^{-2}), cloud droplet number path (CDNP, 10^8 m^{-2}), liquid water path (LWP, g m^{-2}), RH2 (%), and precipitation (mm h^{-1}) to NO_x and VOC emission reductions. Δ_{aer}AOD550 is very small as Δ_{aer}PM2.5 is small. Δ_{aer}SWDOWN is positive under all types and Δ_{aer}CDNP/Δ_{aer}LWP decrease under Convective, Stable, and WarmHumid, both caused by reduced $PM_{2.5}$ after emission reduction. These radiation and cloud perturbations drive the meteorological response already shown in the main text (the sustained Δ_{aer}T2 and Δ_{aer}PBLH increases in Fig. 10 and the Δ_{aer}RH2 decrease in Fig. S15).

We have added Fig. S17 to quantify the impact of aerosol-meteorology feedback on DHP elimination under the full 6×6 matrix of NOx/VOC reduction scenarios. For each scenario we tracked the 1510 baseline DHP samples and calculated DHP elimination, i.e., the fraction of baseline DHP samples removed by emission reductions in each EXP_W_NO(x)_VOC(y) and EXP_WO_NO(x)_VOC(y) (Fig. S17a, b). Fig. S17(c), which is the difference between Figs. S17a and S17b, shows that the feedback enhances DHP elimination under every scenario, by up to 5.9 percentage points, consistent with the enhanced $PM_{2.5}$ response ($\Delta_{aer}PM_{2.5} < 0$) in Section 3.5.

Based on the above reply, we have added Fig. S15 and Fig. S17 to the supplementary file and made the following revisions in the manuscript:

"aerosol feedback" has been revised to "aerosol-meteorology feedback" throughout the whole manuscript.

In the Abstract, the sentence "This study investigated the DHP response to NO_x–VOC emission reductions and the role of aerosol-meteorology feedback in Beijing–

Tianjin–Hebei-Shandong" has been revised to "This study investigated the $PM_{2.5}/O_3$ response to NO_x–VOC precursor perturbations, which jointly control O_3 and nitrate, and the impact of aerosol-meteorology feedback on $PM_{2.5}/O_3$ response in Beijing–Tianjin–Hebei–Shandong in July 2020 using WRF-Chem." to make the full chain explicit.

In the fourth paragraph in Introduction, we have added the full chain definition "Emission controls change aerosol loading, perturbing radiation (ARI) and cloud microphysics (ACI). The resulting adjustments in T2, RH2, and PBLH modulate the $PM_{2.5}$ and O_3 responses, ultimately affecting DHP occurrence.".

In the third paragraph of Section 2.3, the calculation of $\Delta_{aer}PM_{2.5}/O_3$ has been revised to "The response of $PM_{2.5}(O_3)$ concentrations to the emission reductions with and without aerosol-meteorology feedback was obtained by subtracting EXP_W_BASE from EXP_W_NO(x)_VOC(y) and EXP_WO_BASE from EXP_WO_NO(x)_VOC(y), denoted $\Delta PM_{2.5}/O_3$ and $\Delta_{wo}PM_{2.5}/O_3$, respectively. Subtracting $\Delta_{wo}PM_{2.5}/O_3$ from $\Delta PM_{2.5}/O_3$ yielded the impact of aerosol-meteorology feedback on the response of $PM_{2.5}/O_3$ concentration to precursor perturbations, denoted $\Delta_{aer}PM_{2.5}/O_3$.".

In the fourth paragraph of Section 3.5, we have revised the sentence "This indicates that when aerosol feedback is weakened under emission control, the enhanced solar radiation raises surface temperature more, which in turn increases sensible heat flux and promotes boundary layer development." to "This indicates that when aerosol-meteorology feedback is weakened under emission control, the enhanced shortwave downward radiation (SWDOWN) and decreased liquid water path (LWP) (Fig. S15) raise surface temperature."

In the sixth paragraph of Section 3.5, we have added "The aerosol-meteorology feedback also modifies DHP elimination under emission reductions (Fig. S17). Tracking the 1510 baseline DHP samples across the reduction matrix, the feedback enhances DHP elimination by up to 5.9 percentage points, strengthening with deeper NOx reductions but insensitive to VOC reductions, consistent with the enhanced $PM_{2.5}$ response ($\Delta_{aer}PM_{2.5} < 0$).".

Figure S15 Diurnal profiles of the impact of aerosol-meteorology feedback ($\Delta_{aer}X$) on the response of (a) aerosol optical depth at 550 nm (AOD_{550}), (b) shortwave downward radiation (SWDOWN, W m^{-2}), (c) cloud droplet number path (CDNP, 10^8 m^{-2}), (d) liquid water path (LWP, g m^{-2}), (e) RH2 (%), and (f) precipitation (mm h^{-1})

to NO_x and VOC emission reductions under different meteorological condition types. The vertical bars plotted every 3 h indicate the p10–p90 range across emission reduction scenarios.

Figure S17 DHP elimination after NO_x and VOC emission reductions in (a) EXP_W_NO(x)_VOC(y) (%) and (b) EXP_WO_NO(x)_VOC(y) (%), and (c) the impact of aerosol-meteorology feedback on DHP elimination (percentage points). Elimination denotes the fraction of the 1510 baseline DHP samples that are removed by emission reductions.

3. The manuscript defines DHP using daily mean PM2.5 > 35 ug m-3 and MDA8 O3 > 100 ug m-3, while most of the results are reported in ppb. Please include a sensitivity test using alternative DHP thresholds or at least quantify how many grid-days change when the Class I thresholds are perturbed. Because DHP identification drives every subsequent result, the threshold choice should not be presented as self-evidently reliable.

Reply: We thank the reviewer for this important comment.

We have now unified O_3 units throughout the manuscript. For DHP identification, simulated O_3 (in ppb from model output) was converted to μg m^{-3} under the reference state (298.15 K and 1013.25 hPa) specified in China's National Ambient Air Quality Standard (GB 3095–2012 and GB 3095–2026). This ensures that the DHP threshold (MDA8 O_3 > 100 μg m^{-3}) is evaluated consistently with the national standard. All O_3 results in the revised manuscript, including figures (Figs. 3, 4, 6, 7, 8, 10 and Figs. S8, S9, S11, S14) and text, are now presented in μg m^{-3}.

We agree that the robustness of DHP identification to threshold deserves quantitative evaluation, because all subsequent analyses are built on the DHP sample. We have conducted a comprehensive sensitivity test by perturbing the DHP thresholds ($PM_{2.5}$ > 35 μg m^{-3} and MDA8 O_3 > 100 μg m^{-3}) by ±2.5% and ±5% ($PM_{2.5}$: 33.25–36.75 μg m^{-3}; MDA8 O_3: 95–105 μg m^{-3}; a 5 × 5 matrix, 25 combinations). For each combination we repeated the same calculation, including DHP sample identification and meteorological condition classification, and compared the results against the reference threshold ($PM_{2.5}$ > 35 μg m^{-3} and MDA8 O_3 > 100 μg m^{-3}, 1510 DHP samples).

Figure S3 and Table S1 show the sensitivity of DHP identification and meteorological condition classification to the threshold. Figure S3a shows that the DHP sample count is governed almost entirely by the $PM_{2.5}$ threshold (1094–2114, −27.5%

to +40.0%), whereas the five O_3-threshold curves nearly coincide and varying O_3 threshold between 95 and 105 μg m^{-3} changes the count by less than 0.3% as July MDA8 O_3 over BTHS is well above the threshold (overall median 155.4 μg m^{-3}). Figures S3b, S3c and Table S1 show that within each meteorological condition type the concentration statistics change only marginally across all 25 combinations: O_3 means, medians, and quartiles shift by ≤3.2, ≤4.2, and ≤5.9 μg m^{-3}, respectively, while those of $PM_{2.5}$ change by ≤2.3, ≤2.9, and ≤2.7 μg m^{-3}. The contrasts between types persist at every threshold: DryHot and Moderate show the highest O_3, WarmHumid the lowest O_3, Stable the highest $PM_{2.5}$, and DryHot the lowest $PM_{2.5}$. Figure S3d shows that condition proportions remain within narrow bands and the ranking (WarmHumid > Moderate > DryHot > Stable > Convective) is also the same at each threshold. Table S1 complements the figure with per-type statistics. In addition, two-sample Kolmogorov–Smirnov tests detected no significant O_3 distribution change (0/120 pairs, minimum p = 0.078) with Cohen's d < 0.18 (negligible); $PM_{2.5}$ distributions were stable under ±2.5% (p > 0.06, d < 0.17) and changed only marginally under ±5% (d < 0.41). In summary, threshold perturbation mainly rescales the sample count through the $PM_{2.5}$ criterion, whereas the quantities on which our conclusions rest (within-type concentration distributions, between-type contrasts, and the condition ranking) are invariant.

Based on the above analysis, we have added the following discussion to the third paragraph of Section 3.2 and have added Fig. S3 and Table S1 to the supplementary file.

"To test the robustness of DHP identification, Class I thresholds ($PM_{2.5}$ > 35 μg m^{-3} and MDA8 O_3 > 100 μg m^{-3}) were perturbed by ±2.5% and ±5% ($PM_{2.5}$: 33.25–36.75 μg m^{-3}; MDA8 O_3: 95–105 μg m^{-3}; a 5 × 5 matrix, 25 combinations) and the entire filtering and classification process was repeated for each of the 25 threshold combinations. Figure S3 and Table S1 show the sensitivity of DHP identification and meteorological condition classification to the threshold. The DHP count is governed almost entirely by the $PM_{2.5}$ threshold (−27.5% to +40.0%), whereas varying the O_3 threshold between 95 and 105 μg m^{-3} changes the count by less than 0.3%. Within each meteorological condition type, concentration statistics change only marginally across all threshold combinations: O_3 means, medians, and quartiles shift by ≤3.2, ≤4.2, and ≤5.9 μg m^{-3}, respectively, while those of $PM_{2.5}$ change by ≤2.3, ≤2.9, and ≤2.7 μg m^{-3}. The contrasts between types persist at every threshold, with DryHot and Moderate showing the highest O_3, WarmHumid the lowest O_3, Stable the highest $PM_{2.5}$, and

DryHot the lowest $PM_{2.5}$, and the condition ranking (WarmHumid > Moderate > DryHot > Stable > Convective) is preserved at each threshold. In addition, two-sample Kolmogorov–Smirnov tests detected no significant O_3 distribution change (all $p > 0.07$; Cohen's $d < 0.18$), and $PM_{2.5}$ distributions were stable under ±2.5% and changed only marginally under ±5% ($d < 0.41$). Although threshold perturbation mainly rescales the sample count through the $PM_{2.5}$ criterion, the quantities on which our conclusions rest (within-type concentration distributions, between-type contrasts, and the condition ranking) are invariant. The meteorological-condition-specific pollution characteristics underlying our conclusions are therefore robust to the choice of Class I threshold."

Figure S3 Sensitivity of DHP identification and meteorological condition classification to the threshold. (a) DHP sample count as a function of the $PM_{2.5}$ threshold; each line denotes one MDA8 O_3 threshold (95–105 μg m^{-3}), with the black solid line indicating the reference. (b) $PM_{2.5}$ and (c) MDA8 O_3 medians within each meteorological condition type: dots denote the reference threshold and vertical bars the range across all 25 combinations. (d) Meteorological condition proportions as a function of the $PM_{2.5}$ threshold with the O_3 threshold fixed at 100 μg m^{-3}; the dashed line marks the reference $PM_{2.5}$ threshold.

Table S1 Statistics of the DHP samples at the reference thresholds ($PM_{2.5} > 35$, MDA8 O_3 > 100 μg m^{-3}) and their ranges across the 25 perturbed threshold combinations. Medians are in μg m^{-3}.

4. EXP_W and EXP_WO differ simultaneously in aerosol-radiation interaction, aerosol-cloud interaction, cloud microphysics, radiation, and potentially the meteorological evolution. Setting cloud droplet number concentration to 10 cm-3 over continental BTHS is a strong intervention and is not equivalent to simply "turning off ACI." The difference between the two experiments therefore represents a combined response to the chosen aerosol-coupling perturbation, not an isolated, uniquely attributable aerosol-meteorology feedback. The revised manuscript should also quantify changes in other cloud properties including cloud water, CDNC, and key meteorological fields between EXP_W and EXP_WO, and should explain whether the fixed-CDNC configuration changes wet removal or cloud lifetime.

Reply: We thank the reviewer for this insightful comment.

We agree that prescribing a fixed CDNC of 10 cm^{-3} is not equivalent to simply "turning off ACI". EXP_WO differs from EXP_W simultaneously in ARI, ACI, cloud

microphysics, radiation, and the subsequent meteorological evolution, so the difference between EXP_W and EXP_WO represents the combined response to this aerosol-coupling perturbation rather than an isolated, uniquely attributable feedback. The description of EXP_WO has been revised.

We have added Fig. S16 to quantify the cloud properties and key meteorological fields (AOD550, SWDOWN, CDNP, LWP, T2, RH2, PBLH, and precipitation) between EXP_W and EXP_WO. The feedback reduces SWDOWN, increases CDNP and LWP, lowers T2 and PBLH, and raises RH2, with the largest responses under Stable. Since the DHP identification excludes days with more than 3 h of precipitation, the precipitation differences between the two experiments are very small, indicating that the fixed-CDNC-induced changes in wet removal and cloud lifetime have limited influence on the diagnosed responses.

Based on the above analysis, we have added Fig. S16 to the supplementary file and have revised the manuscript as follows.

In the first paragraph of Section 2.3, The description of EXP_WO has been revised to "The first experiment set (EXP_W) accounted for the full interaction between aerosols and meteorology including both ARI and ACI. The second experiment set (EXP_WO) excluded both ARI and ACI. ARI was disabled by removing aerosol effects on radiation in the RRTMG scheme, and ACI was disabled by prescribing a fixed cloud droplet number concentration (CDNC) of 10 cm^{-3} in the Morrison double-moment cloud microphysics scheme (Zhao et al., 2017; Zhang et al., 2018; Xiong et al., 2022; Zhao et al., 2026b), a value consistent with satellite observations of clean marine liquid clouds (Bennartz, 2007; Zeng et al., 2014). EXP_WO differs from EXP_W simultaneously in both ARI and ACI, and the subsequent meteorological evolution, so the diagnosed feedback represents the combined response to this perturbation."

In the sixth paragraph of Section 3.5, we have added "The results of Fig. 10 are also supported by Fig. S16, which quantifies the differences in cloud properties and key meteorological fields between EXP_W and EXP_WO. The feedback reduces SWDOWN, increases cloud droplet number path (CDNP) and LWP, lowers T2 and PBLH, and raises RH2, with the largest responses under Stable. Since the DHP identification excludes days with more than 3 h of precipitation, the precipitation differences between the two experiments are very small, indicating that the fixed-CDNC-induced changes in wet removal and cloud lifetime have limited influence on the diagnosed responses.".

Figure S16 Diurnal profiles of the difference in (a) O_3 (μg m^{-3}), (b) $PM_{2.5}$ (μg m^{-3}), (c) AOD_{550}, (d) SWDOWN (W m^{-2}), (e) CDNP (10^8 m^{-2}), (f) LWP (g m^{-2}), (g) T2 (°C), (h) RH2 (%), (i) PBLH (km), and (j) precipitation (mm h^{-1}) between EXP_W and EXP_WO under different meteorological condition types.

5. The analysis covers 1-30 July 2020 only. July 2020 may contain unusual emissions, circulation, humidity, and post-lockdown conditions, and the five classes may not represent DHP in other seasons or years. The limitations section acknowledges this but the abstract, conclusions, and management implications remain broad.

Reply: We thank the reviewer for this insightful comment.

July 2020 was chosen as a representative summertime testbed as it falls in the peak season of DHP in northern China (Wang et al., 2025a) and provides sufficient DHP samples (1510 grid-cell days). Regarding emissions, China's COVID-19 restrictions had been largely lifted by July 2020, and the simulations use the MEIC 2020 inventory; more importantly, our experimental design applies relative NOx/VOC reductions (0–50%) to a fixed baseline, so the conclusions concern the sensitivity structure of the chemical regime and the aerosol feedback rather than the absolute emission level of this particular month. The favorable meteorological regimes identified here (warm, moderately humid, shallow boundary layer) are consistent with those documented in multi-year studies of co-occurring $PM_{2.5}/O_3$ pollution over the North China Plain (Ma et al., 2023; Luo et al., 2022; Dai et al., 2024; Liu et al., 2026), indicating that the five regimes are not unique to 2020.

Following the reviewer's suggestion, we have revised the title, some sentences in the abstract, conclusion, and management implications with case-study specifics.

We have added "in July 2020" to the title to explicitly state the study period, so that the single-month scope of this case study is clear.

In the Abstract, the sentence "These results demonstrate that meteorological condition-specific strategies, particularly coordinated NOx–VOC control under stagnant and humid conditions, are essential for effective DHP mitigation." has been revised to "These results demonstrate that meteorological condition-specific strategies, particularly coordinated NOx–VOC control under stagnant and humid summer conditions, are essential for effective DHP mitigation."

In the sixth paragraph of Conclusion, the sentence "First, the analysis covers a single month (July 2020); while the five meteorological conditions types are physically

distinct, they may not capture the full spectrum of conditions conducive to DHP in other seasons or years." has been revised to "First, as the analysis covers a single month (July 2020), the five meteorological condition types, though physically distinct and consistent with multi-year studies, may not represent the frequency distribution and NOx–VOC sensitivity of such conditions in other seasons or years."

In the seventh paragraph of Conclusion, we have added the following sentence to the management implications: "It is noted that these implications are drawn from a single-month case study and should be re-examined for other seasons and years before operational application."

Minor comments:

1. Use one regional abbreviation consistently. The manuscript alternates among BTHS, BTH, SD, and "Beijing-Tianjin-Hebei-Shandong"; define the spatial domain and the observational subregions once and use them consistently.

Reply: We thank the reviewer for this important suggestion.

We have now standardized the regional abbreviation throughout the manuscript. Currently, only "Beijing–Tianjin–Hebei–Shandong (BTHS)", "Beijing–Tianjin–Hebei", and "Shandong" are used, of which "Beijing–Tianjin–Hebei" and "Shandong" are only used when presenting model validation results for the sub-regions in Section 3.1. Elsewhere in the manuscript, we now consistently use BTHS when referring to the entire study domain, and we have removed all inconsistent uses. In addition, we have revised Fig. 1, Fig. 2, and Table 2 to ensure consistent regional labeling.

2. Correct the unit and wording errors in the text and captions, including "45% to 80%°C" in Sect. 3.2, "Figrure 7," "effects" where "affects" is intended, and the double-minus notation "--0.12" in the discussion of Fig. 10.

Reply: We thank the reviewer for pointing out these errors.

All errors have been corrected: "45% to 80%°C" is now "45% to 80%", "Figrure 7" is corrected to "Figure 7", "effects" is changed to "affects" where a verb is required, and "--0.12" is replaced with "−0.12".

3. Standardize O3 units throughout. Table 2 reports O3 in ug m-3, while the text and most figures use ppb.

Reply: We thank the reviewer for pointing out this inconsistency.

We have now unified O_3 units throughout the manuscript. All O_3 results in the revised manuscript, including figures (Figs. 3, 4, 6, 7, 8, 10 and Figs. S8, S9, S11, S14) and text, are now presented in μg m^{-3}.

4. The captions of Fig. S7 use labels such as Stable_Inv, Neutral_WarmHumid, Neutral_DryHot, and Neutral_Cool, which do not match the five classes used in the main text.

Reply: We thank the reviewer for pointing this out.

We have updated the Fig. S7 (now Fig. S8) captions to use the same five weather-type labels as in the main text as follows: "**Figure S8** MDA8 O_3 and $PM_{2.5}$ response to NO_x and VOC reductions for DHP samples classified into different meteorological condition types. The absolute differences in MDA8 O_3 between each reduction scenario and the baseline scenario (ΔMDA8 O_3, μg m^{-3}) for (a) Convective, (b) Stable, (c) WarmHumid, (d) DryHot, and (e) Moderate, and the absolute differences in $PM_{2.5}$ between each reduction scenario and the baseline (Δ$PM_{2.5}$, μg m^{-3}) for (f) Convective, (g) Stable, (h) WarmHumid, (i) DryHot, and (j) Moderate. Rows and columns denote NO_x and VOC reduction levels (%)."

5. In Sect. 3.2, the statement that DHP O3 is "consistently above the all samples median" is inconsistent with the reported WarmHumid median of about 76 ppb versus an all-sample median near 79 ppb. Please revise the statement.

Reply: We thank the reviewer for pointing out this inconsistency.

The dashed reference line representing the all-sample median was drawn incorrectly: it showed the median of all samples under a single meteorological type rather than the overall median of all samples. The boxplots and other elements were unaffected. We have corrected the calculation and redrawn the dashed line in all affected figures (Figures 4, S5, S6, and S7). The corresponding descriptions in the text have also been updated to "MDA8 O_3 remains elevated in all types (medians ~149–164 μg m^{-3}), consistently above the median of all samples (~140 μg m^{-3})."

6. Please proofread the manuscript for grammar and spacing, for example "conditions types," "which is attribute," missing spaces around parentheses, and inconsistent use of hyphen/en dash in compound terms.

Reply: We thank the reviewer for this careful proofreading.

We have thoroughly checked the entire manuscript and corrected all identified grammar and spacing errors, including “conditions types” (now “condition types”), “which is attribute” (now “which is attributed”), missing spaces around parentheses, and inconsistent hyphen/en dash usage in compound terms.

Response to the general comments in the review summary:

We thank the reviewer for the careful overall assessment. Two concerns raised in the summary paragraph (the model biases and the effective sample size/uncertainty) were not listed as numbered comments, and we address them here to make the revision complete.

(1) For model biases, we acknowledge the PBLH, wind speed, and nighttime O_3 biases quantified in Section 3.1 and Table 2. We note that these biases have limited influence on the central analysis for three reasons. First, the central analysis is based on relative responses of each perturbation scenario to the baseline experiment. Second, the meteorological classification is defined by the percentiles of the model’s own PBLH distribution and temperature/humidity thresholds, making the classification internally consistent and insensitive to the absolute PBLH bias. Third, although the nighttime O_3 low bias affects the simulated MDA8 values, the DHP screening applies a uniform threshold across all meteorological condition types, so the type-to-type contrasts are unaffected.

To make this explicit, we have added the following sentence to the third paragraph of Section 3.1: “We note that the central analysis of this study is based on relative responses and internally consistent classification, so the model biases documented above have limited influence on the type-to-type contrasts.”

(2) For effective sample size and uncertainty, we agree that the 1510 DHP samples are not fully independent because neighboring grids are spatially correlated and pollution persists from day to day; the effective sample size is therefore smaller than the nominal count. The main conclusions are nevertheless robust to this dependence. Specifically, for $PM_{2.5}$, the between-type contrast (Fig. 4a) remains robust even if the effective sample size is only a fraction of the nominal count, given the 87–604 samples per type. For O_3, the key conclusion is its persistent elevation under all five types. For the aerosol-meteorology feedback diagnostics, the p10–p90 ranges across the 35 reduction scenarios (Figs. 10 and S13) further show that the sign of the diagnosed effects is robust to the reduction pathway. The robustness of the sample statistics to the

DHP threshold choice is additionally quantified in Section 3.2 (Fig. S3 and Table S1).

To acknowledge this limitation explicitly, we have added the following sentence to the sixth paragraph of Conclusions: “The DHP samples are also not fully independent owing to spatial and temporal autocorrelation, so the effective sample size may be smaller than the nominal count; the main conclusions, however, rest on large between-type contrasts and the persistent elevation of O_3 under all types, which are robust to this dependence.”