## Response to Referee #1's comments

This is a very solid and scientifically valuable piece of work. The authors conducted a systematic, year-long observational analysis of horizontally oriented ice crystals (HOICs) using a unique ground-based combination of dual-angle polarization lidars and a cloud radar in Beijing. For the first time, the study provides long-term statistical constraints on the occurrence frequency, macrophysical characteristics, environmental conditions, and relationships of HOICs with supercooled liquid water clouds (SWCs) and turbulence, filling a critical gap in this field. The data are of high quality, the analysis methodology is rigorous, and the results are of significant reference value for improving ice-cloud microphysics parameterization schemes in climate models. The paper is well structured and the arguments are well supported, but some details still need improvement. Therefore, a minor revision is recommended before acceptance.

**Reply:** We sincerely thank Referee #1 for reviewing our paper and providing constructive comments for improvement. Responses to these comments are provided below. In this author's comment, we reply (in blue font) to Referee #1's comments (in black font). The relevant part in the revised version is in green font.

1. Is Figure 5(a) based on statistics over all temperatures? The fraction of water clouds in 2022 is only 5.2% – could this be double-checked?

**Reply:** Thank you for the comment. Figure 5(a) is based on statistics over all temperatures. We have double-checked the calculation and confirmed that warm liquid-water clouds account for 5.2% of all identified cloud range bins in 2022. It should be emphasized that this percentage represents a **range-bin-based fraction (5 min × 15 m)** rather than the fraction of time during which liquid-water clouds occurred.

Stratiform liquid-water clouds generally have a smaller vertical extent than ice clouds, and optically thick liquid-water clouds may also be underrepresented because strong lidar attenuation limits the penetration of the laser beam. In contrast, vertically extended ice clouds contribute a much larger number of cloud range bins for a given observation period. Therefore, the relatively small fraction of warm liquid-water cloud range bins does not imply that liquid-water clouds occurred during only 5.2% of the cloudy periods.

For comparison, in our previous paper (Wu et al., 2026), the fraction of water clouds is 5.3 % using rigorous cloud mask.

2. For Figure 6 (rigorous cloud mask) and Figure 7 (relaxed cloud mask), the numbers of statistical cases are 735 and 693 respectively, with a difference of more than 5%. Could the corresponding reasons be provided?

**Reply:** Thank you for the comment. Although the relaxed cloud mask identifies more HOIC range bins than the rigorous cloud mask, many of these additional bins fill temporal or vertical gaps between HOIC fragments that are treated as separate events under the rigorous mask. According to our event-grouping criteria ($\Delta t$ < 60 min, $\Delta h$ < 2 km), these newly connected HOIC fragments are therefore merged into the same event. Consequently, the relaxed cloud mask yields fewer but more complete individual HOIC events (693) than the rigorous cloud mask (735). We have added the corresponding explanation to the revised manuscript (Line 469–472):

"Although the relaxed cloud mask identifies more HOIC range bins, many of the additional bins bridge temporal or vertical gaps between HOIC fragments that are treated as separate events under the rigorous mask. Under the event-grouping criteria ($\Delta t < 60$ min and $\Delta h < 2$ km), these fragments are therefore merged into fewer but more complete HOIC events, reducing the number of identified events from 735 to 693."

3. The paper cites previous studies multiple times, but as an independent paper for publication, it should appropriately explain those previous studies. For example, in the sentence "Moreover, HOICs produce much stronger backscatter in the zenith-pointing lidar than in the off-zenith-pointing lidar, making cloud-layer detection more easily triggered in the zenith observations (Zhao et al., 2014; Wu et al., 2025)," the physical mechanism should be articulated within this paper.

**Reply:** Thank you for this important comment. We agree that the physical mechanism should be explained explicitly rather than only referred to through citations. We have therefore revised the relevant part of Sect. 2.2 to clarify the angle-dependent specular reflection of HOICs and its influence on cloud-layer detection. The relevant passage in Sect. 2.2 has been revised as follows:

"Moreover, HOICs can produce mirror-like specular reflection. For the zenith-pointing lidar, the incident laser beam is nearly perpendicular to the main facets of the HOICs, so a strong specular reflection component is returned to the receiver. At 15° off zenith, the incident laser beam is no longer close to normal incidence, and most of the specular reflection component is reflected away from the receiver, resulting in much weaker backscatter (Wu et al., 2025). Because the cloud-layer detection algorithm identifies cloud bins when the backscatter-related quantity exceeds a predefined threshold, HOIC-containing cloud bins are more readily detected in the zenith observations (Zhao et al., 2014)."

4. The paper states: "both confirming a statistically significant difference in their overall EDR distributions ($p \ll 0.001$). Interestingly, the calculated Cliff's Delta yields a small overall effect size ($d = -0.060$)." The formulas for p and d should be provided.

**Reply:** Thank you for this constructive suggestion. We fully agree that providing explicit mathematical formulas for both the statistical hypothesis test (p-value calculation via the Mann–Whitney U test and Kolmogorov–Smirnov test) and the effect size metric (Cliff's Delta, *d*) enhances the clarity and reproducibility of our methodology. We have now added the mathematical definitions and formulas for *p* and *d* in Sect. 2.5:

"Let $X = \{x_i\}_{i=1}^{m}$ and $Y = \{y_j\}_{j=1}^{n}$ denote the cloud-property samples for HOICs and ROICs, respectively. The Mann–Whitney U statistic corresponding to the HOIC sample was calculated from the pooled ranks as:

$$U_{\text{obs}} = R_X - \frac{m(m+1)}{2} \quad (1)$$

where $R_X$ is the sum of the pooled average ranks assigned to the $m$ HOIC observations. A one-sided alternative, assuming that the HOIC values are stochastically smaller than the ROIC values, was used, and its *p*-value was defined as

$$p_{\text{MW}} = \Pr_{H_0}(U \le U_{\text{obs}}) \quad (2)$$

where $U_{\text{obs}}$ is the observed Mann–Whitney statistic, $U$ denotes the corresponding random statistic under the null hypothesis $H_0$, and $\Pr_{H_0}$ denotes probability under the null hypothesis $H_0$ that the HOIC and ROIC samples originate from the same underlying distribution. Given the large sample sizes, the Mann–Whitney *p*-value was evaluated using the asymptotic implementation in SciPy (scipy.stats.mannwhitneyu; Virtanen et al., 2020), with correction for ties and the default continuity correction.

For the two-sided two-sample Kolmogorov–Smirnov test, the observed statistic was defined as:

$$D_{\text{obs}} = \sup_z |F_m(z) - G_n(z)| \quad (3)$$

with the corresponding *p*-value

$$p_{\text{KS}} = \Pr_{H_0}\left(D_{m,n} \ge D_{\text{obs}}\right) \quad (4)$$

where $F_m$ and $G_n$ are the empirical cumulative distribution functions of the HOIC and ROIC samples, respectively; $\sup_z$ denotes the supremum over all possible $z$; $D_{m,n}$ denotes the random two-sample Kolmogorov–Smirnov statistic for sample sizes $m$ and $n$ under $H_0$; and $H_0$ denotes that the two underlying distributions are identical. The *p*-value was evaluated via the asymptotic Kolmogorov–Smirnov test in SciPy (scipy.stats.ks_2samp; Virtanen et al., 2020)."

We further added:

"Cliff's Delta was calculated directly from the Mann–Whitney statistic returned for the HOIC sample:

$$d = \frac{2U_{\text{obs}}}{mn} - 1 \quad (5)$$

where *m* and *n* are the HOIC and ROIC sample sizes, respectively. "

We also corrected several instances in which hyphens were incorrectly used instead of en dashes.

We have also corrected the reference for the two-sample Kolmogorov–Smirnov test. The previously cited goodness-of-fit paper by Massey (1951) has been replaced with the paper specifically addressing the maximum deviation between two empirical cumulative distribution functions: Massey, F. J., Jr. (1951), "The distribution of the maximum deviation between two sample cumulative step functions."

**Reference:**

Wu, Z., Li, C., He, Y., Baars, H., and Seifert, P.: Horizontally Oriented Ice Crystals Observed with the Combination of Zenith and 15-degree off-Zenith pointing Polarization Lidar over Beijing

(116.3°E40.0°N), China, EPJ Web Conf., 362, 02021, https://doi.org/10.1051/epjconf/202636202021, 2026.

Massey, F. J., Jr.: The distribution of the maximum deviation between two sample cumulative step functions, Ann. Math. Statist., 22, 125–128, https://doi.org/10.1214/aoms/1177729703, 1951.