<u>Response to Community Comments 2</u>

Thank you for your comment on our manuscript. The authors have some skepticism regarding your concerns, as you refer to published uncertainties in the level 3 data, but the level 2 data are used in the current study and thus the level 3 uncertainties may not be applicable here. For example, you state that the error in the level 3 maxima is 5-7 m/s, but the published uncertainty in the Vaisala wind speeds is only 0.15 m/s. The authors do acknowledge that the "real" uncertainty may be higher but likely are not on the order of 5-7 m/s. A comparison of radiosonde winds at 10 m to the met tower winds at 10 m from MOSAiC measurements during October – April shows an RMSE of 1.17 m/s. Additionally, the RMSE between the ship's wind measurement (29 m) to the radiosonde wind speed at 30 m is 1.27 m/s. A comparison of the DataHawk2 UAS and radiosonde measurements in Hamilton et al. (2023) reveals a difference of <1 m/s based on the 95% confidence intervals of observations from both platforms. To ultimately determine if wind uncertainties may affect the results in the current analysis, we have repeated the analysis per your suggestions (below). In the end we have determined that the original methods (using a threshold of 2 m/s and no vertical averaging) give the most accurate climatology (see Sect. 1 of response) and are relatively unaffected by the errors you refer to (see Sect 2 of response).

Hamilton, J., de Boer, G., Doddi, A., and Lawrence, D. A.: The DataHawk2 uncrewed aircraft system for atmospheric research, Atmos. Meas. Tech., 15, 6789–6806, https://doi.org/10.5194/amt-15-6789-2022, 2022.

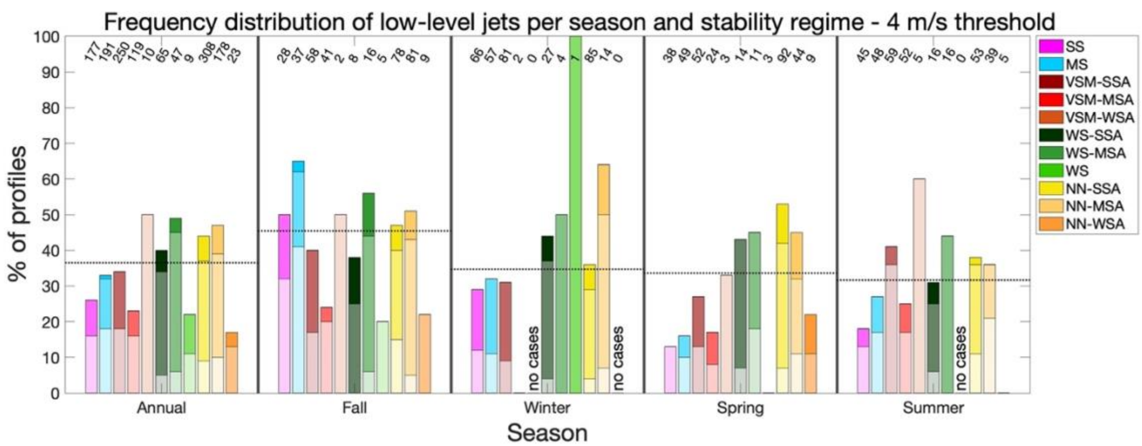
Section 1: Changing the criterion for identifying a wind speed maximum as an LLJ.

First, we repeated the routines regarding LLJs and have reproduced the corresponding figures, now using a threshold of 4 m/s instead of 2 m/s. The figures are shown below. The authors found that changing the threshold this drastically yields an entirely different LLJ climatology that is less likely to be true (see reasons below). Given these results, the authors disagree that applying a 4 m/s threshold is an appropriate approach.

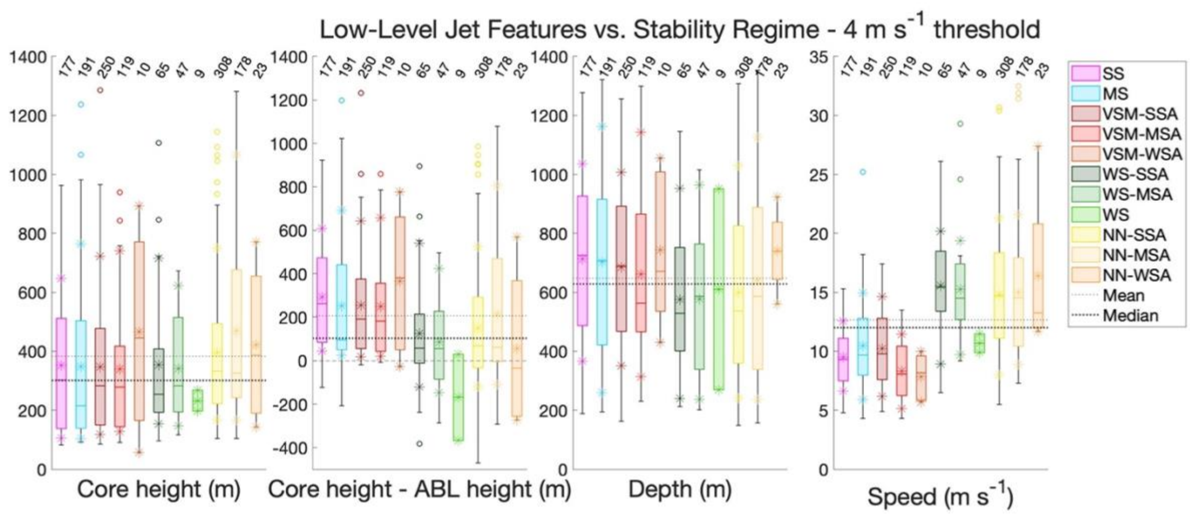
1. Using this threshold of 4 m/s (and not using the 25% criterion) now gives an annual LLJ frequency of 36.5%. This is now a much lower frequency than was found in by Lopez-Garcia et al. (2022). The discrepancy between our LLJ frequency and that of by Lopez-Garcia et al. (2022) in the submitted draft of this manuscript can be attributed to the fact that we did not use the 25% criterion, whereas they did. Otherwise, all other methods are consistent between the two papers, which was done intentionally through communications between the authors of the current paper and the authors of by Lopez-Garcia et al. (2022). Depending on a reader's purpose in wanting to know about LLJ frequency and characteristics, they can then draw from either the results of the current paper or those of by Lopez-Garcia et al. (2022). However, if we change our methods further by now applying a 4 m/s threshold instead of 2 m/s, this is in direct conflict with Lopez-Garcia et al. (2022) which, as both studies use the same dataset, would be even more confusing to a reader.
2. The only LLJ threshold that appears in previous literature, as far as the authors have found, is the threshold of 2 m/s. In many other studies, this threshold has been applied to radiosonde datasets which likely have uncertainties comparable to those of the level 2 data used in the current study. The authors do not think it makes sense to challenge those methods used by so

many prior studies, as this would make the results of the current study incomparable with those of prior studies.

3. An LLJ frequency of 36.5% is now very low compared to that found in previous work in polar sea ice regions. In our response to your previous comment, we shared several citations which reveal an LLJ frequency in polar sea ice regions closer to the 76% shared in the submitted manuscript for the current study. Thus, a threshold of 4 m/s produces results inconsistent with previous work, so we can be confident that it is likely missing many true LLJ events.
4. One of the primary purposes of the current study is to relate LLJ characteristics to stability regime. The trends presented in the box and whisker plot when a threshold of 4 m/s is used do not differ from the results when a threshold of 2 m/s is used. The primary difference is that the mean and median LLJ speeds are slightly higher when a threshold of 4 m/s is used. This suggests that using a threshold of 4 m/s misses a lot of LLJ events which have slower wind speeds, but are still important in their interaction with near-surface stability (as is the focus of the current study).



As in Supplementary Fig. S4, but when a threshold of 4 m/s is used.



As in Fig. 8, but when a threshold of 4 m/s is used.

Section 2: Applying vertical averaging.

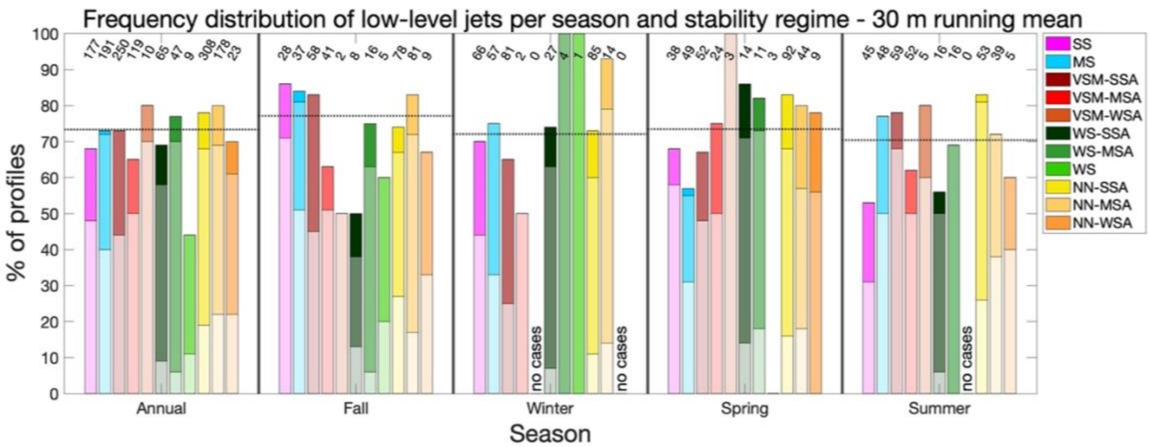
We have also repeated the routines regarding LLJs and have reproduced the corresponding figures, using a running mean of 30 m and 60 m, which should reduce the random error by 60% and 71% respectively (see the math below). Since the claimed error in the wind speed maximum could be as high as 5-7 m/s, a reduction of 60% of the low end of this range (5 m/s) and a reduction of 71% of the high end of this range (7 m/s) would each be 2 m/s, thus bringing the threshold of 2 m/s for a wind speed maximum to be identified as an LLJ within the claimed error.

The error for one data point is:
1/(sqrt(1))*ERROR =1*ERROR

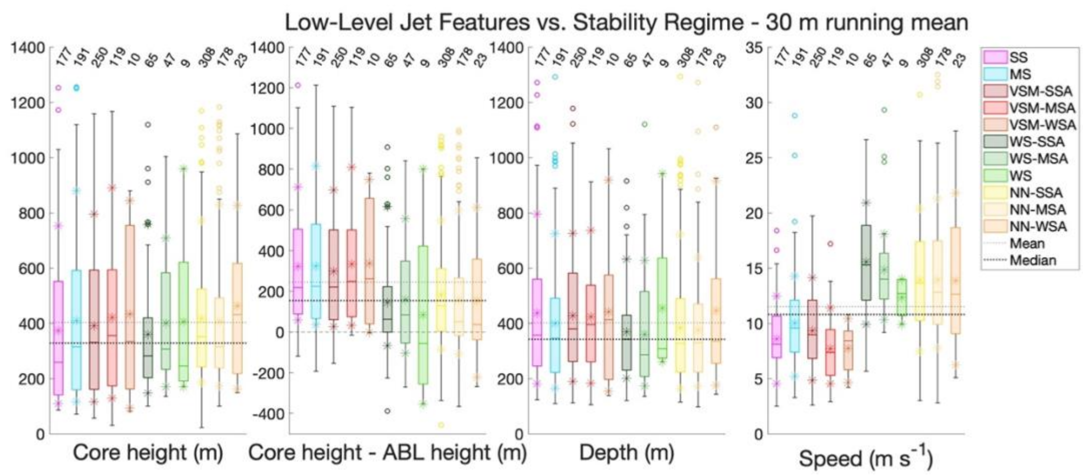
Averaging over 6 data points (i.e. 30 meters vertically with 5 m resolution data), the error for one data point would then be:
1/(sqrt(6))*ERROR=0.4*ERROR, thus reducing the error by 60%
0.4 * 5 m/s = 2 m/s

Averaging over 12 data points (i.e. 60 meters vertically with 5 m resolution data), the error for one data point would then be:
1/(sqrt(12))*ERROR=0.29*ERROR. thus reducing the error by 71%
0.29 * 7 m/s = 2 m/s

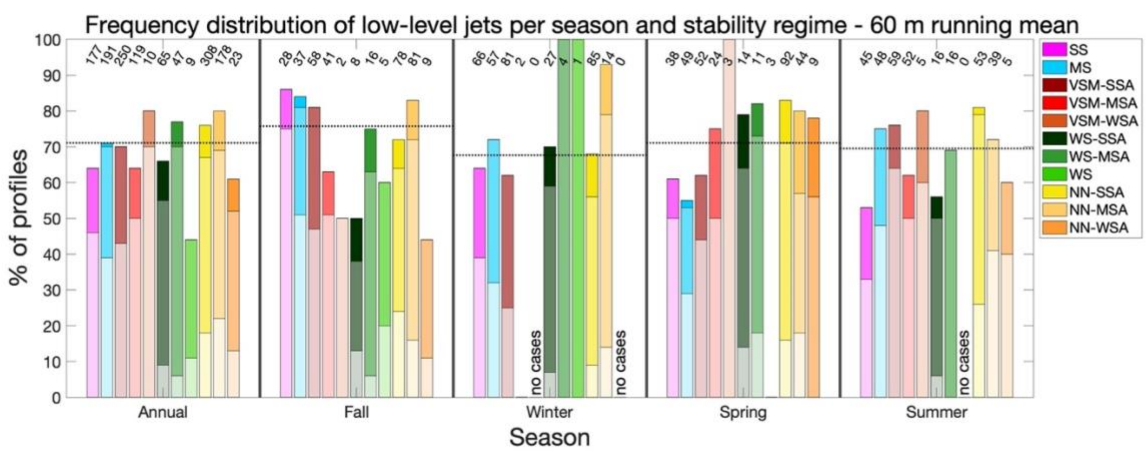
Corresponding figures are shown below. When a 30 m running mean is used, the annual frequency of LLJs is reduced by less than 3%, to 73.4%. When significance testing is conducted, it is found that this frequency is not significantly different at the 95% confidence level from the frequency when no running mean is used. We have added a sentence to the manuscript addressing the aforementioned result. When a 60 m running mean is used, the annual frequency of LLJs is reduced by about 5%, to 71%. This time when significance testing is conducted, it is found that this frequency is significantly different at the 95% confidence level from the frequency when no running mean is used. However, an actual uncertainty of 7 m/s for the maximum in the wind speed from the level 2 data is unlikely to be true, especially given the published uncertainty in the Vaisala wind speeds of 0.15 m/s. Regardless, a reduction of LLJ frequency from 76% to 71% does not change the story that LLJs are commonly occurring in the central Arctic. Additionally, once again applying a running mean does not change the relationships between LLJ characteristics and stability regimes shown in the submitted manuscript, which is the focus of the current study.



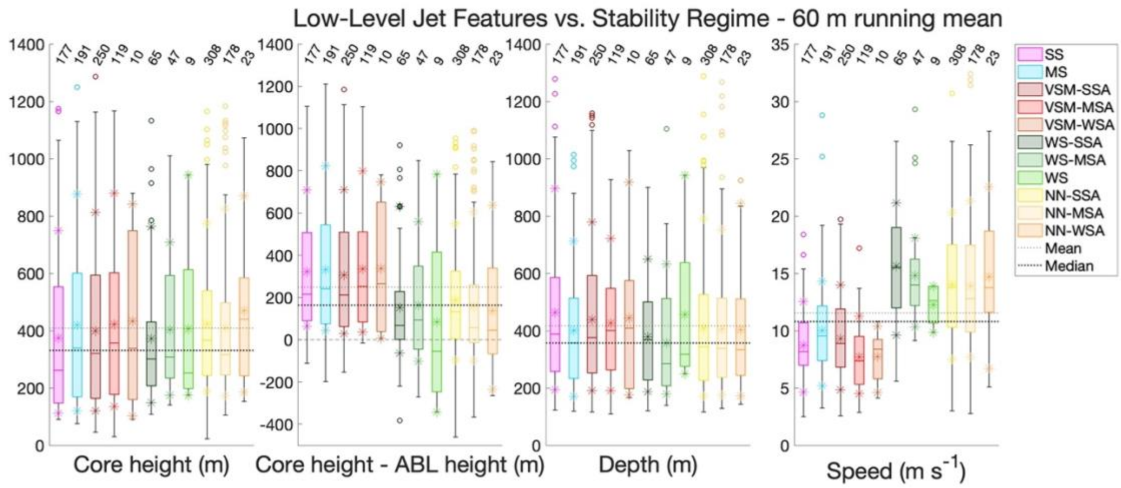
As in Supplementary Fig. S4, but when a 30 m running mean is used.



As in Fig. 8, but when a 30 m running mean is used.



As in Supplementary Fig. S4, but when a 60 m running mean is used.



As in Fig. 8, but when a 60 m running mean is used.