The authors are grateful to the editor and all reviewers for their time and energy in providing helpful comments that have improved the manuscript. In our revised paper, we further re-checked all revisions and performed grammatical corrections to help readers understand our manuscript easier.

In this document, reviewer' comments have been addressed point by point. Referee comments are shown in black italics and author responses are shown in blue regular text and revised version of the manuscript is shown in green text.

Reviewer #1:

Major comments:

The manuscript presents the Integrated Multi-source Polar Meso-Cyclone Tracks (IMPMCT) dataset based on both ERA5 reanalysis and remote sensing data during winter in the Nordic Sea, demonstrates clearly the workflow of this method, and compares the results with existing manually identified and reanalysis-based track datasets. There remains a clear need for establishing a more comprehensive tracking dataset capable of capturing PMCs throughout their lifecycle due to their impacts on human activities and regional climate change. The manuscript is generally well-organized, and the figures effectively communicate the results while being concise. However, there are a few aspects where the presentation could be improved. The detailed comments are listed below, and I encourage the authors to make the necessary adjustments to improve the study.

1. The present study utilized a series of datasets, including ERA5 reanalysis, AVHRR data andQuikSCAT/ASCAT data, which have different spatial and temporal resolutions, and these data are stored with different projections/grids. How are these multi-source datasets treated in the cyclone tracking algorithm to maintain consistency? Please clarify.

Re: Thank you very much for your inquiry. We fully understand your concern about the data matching method and the accuracy of the tracking algorithm. Issues such as spatial-temporal resolution and potential representativeness errors are indeed key issues that must be handled carefully in dataset establishment. We provide the following detailed explanations:

(1) data matching

- **ERA5-AVHRR Matching**: Vortex centers from ERA5 (hourly vorticity fields) were matched to AVHRR cloud features within a 1-hour window and 250-km radius. AVHRR data validated genuine cyclone evolution. Trajectories were excluded if AVHRR temporal resolution was insufficient to confirm evolution or if the average matching distance exceeded 150 km.
- **QuikSCAT/ASCAT**: Wind data supplemented cyclone attributes but did not drive identification. Matches to AVHRR were constrained to a 30-minute window. Scan timestamps are provided for error assessment.

(2) data grids:

We used a VCI(Vortex-Centered Infrared) grid, which is a conformal projection grid. This grid has mutually perpendicular meridians and parallels, with shape invariance under translation, and local equidistant characteristics. It preserves local isotropy and enables consistent spatial calculations (see **[Line 353-363]** for details).

2. Line 164: ERA5 data. How accurate are the ERA-5 fields used in the analysis of the Nordic Sea? What are the known biases? As the authors did not repeat their method with other reanalysis datasets to test the robustness of their results, I would suggest

declaring the known biases of ERA5 in this part.

Re: Your reminder is very important, which helps to improve the rigor of the study. We have added descriptions about the quality of ERA5 regarding meteorological elements related to polar mesocyclones in the Nordic Sea in the revised version, see **Line 169-179**.

This additional ERA5 data description is now described in the revised version of the manuscript:

ERA5 reanalysis dataset demonstrates robust performance in representing meteorological fields over the Nordic Seas, such as sea level pressure, air temperature, and humidity (Graham et al., 2019; Moreno-Ibáñez et al., 2023; Yao et al., 2021). Most notably, its effective characterization of cold air outbreaks has been proven to correlate closely with the timing and location of PLs (Meyer et al., 2021). However, beyond the previously mentioned underestimation of near-surface strong winds in Section 1, Wang et al. (2019) found ERA5 data exhibits a warm bias over Arctic sea ice during winter and spring, which makes it difficult to accurately simulate the frequently occurring strongly stable boundary layers prevalent in winter and early spring. Consequently, the intensity of PMCs near the sea ice edge might be overestimated. Nevertheless, more accurate total precipitation and snowfall data in ERA5 (Wang et al., 2019) significantly benefits the representation of enhanced latent heat release mechanisms associated with PLs (Moreno-Ibáñez et al., 2021).

3. Line 262: To maximize the inclusion of potential PMCs, we implement more lenient vortex detection criteria compared to Stoll et al. (2021). The selected criteria seem to be very subjective. Importantly, how sensitive are the results to subjective criteria such as the “vorticity peak threshold”, “isolated vortex threshold”? Have the authors conducted sensitivity tests, and what metrics were used to evaluate the robustness of the results? Please include this.

Re: Your suggestion is very important. Following your advice, we have deleted the statement that directly adopts lenient criteria to avoid confusing readers. Meanwhile, we have added a subsection "3.1.3 Sensitivity experiments of vortex identification parameters", in which we supplemented two groups of sensitivity experiments on vortex identification parameters. We also calculated the matching rates of vortex tracks obtained from different parameter sets with other PL lists.

Through the experiments, we found that:

- Lowering the vorticity peak threshold (ζ_{max0}) increased detection of weak vortices (lifespan +3 hrs) and nearly doubled capture of moderately weak systems. **[Line 316-322]**
- Reducing the isolation threshold (γ) improved sensitivity to splitting events but shortened mean vortex lifespan by ~2 hrs due to increased transient sub-vortices. **[Line 323-328]**
- Experiment a was chosen to maximize weak-PMC inclusion and validation against PL datasets (Table S2) shows the lenient-threshold vortex tracks consistently yield higher matching rate. **[Line 527-532]**

This additional **Sensitivity experiments** is now described in the revised version of the manuscript:

To evaluate the sensitivity of vortex identification parameters, we conducted three sensitivity experiments with the following configurations, each designed to test the impact of varying key thresholds ζ_{max0} (ζ_{min0}) and γ on vortex detection:

1) Experiment a (lenient thresholds): ζ_{max0} = 1.2×10^{-4} s^{-1}, ζ_{min0} = 1.0×10^{-4} s^{-1}, γ = 0.15;

2) Experiment b (intermediate thresholds): ζ_{max0} = 1.2×10^{-4} s^{-1}, ζ_{min0} = 1.0×10^{-4} s^{-1}, γ = 0.25;
3) Experiment c (strict thresholds, following Stoll et al. 2021): ζ_{max0}= 1.5×10^{-4} s^{-1}, ζ_{min0} = 1.2×10^{-4} s^{-1}, γ = 0.25

The influence of threshold variations on vortex detection characteristics was systematically evaluated by analyzing differences in the number of identified vortex tracks, their lifespans, and their vorticity across the three experiments. As shown in Fig. 7, threshold adjustments predominantly affected vortices exhibiting maximum vorticity (ζ_{trmax}) less than 2×10^{-4} s^{-1}, with distinct impacts observed for changes in ζ_{max0} versus γ. The principal findings are:

First, focusing on the impact of ζ_{max0} (by comparing Experiment b, which uses a lenient ζ_{max0}, with Experiment c, which uses a strict ζ_{max0}), we found that the lenient threshold in Experiment b captured an additional 8,077 weak-vorticity tracks (with ζ_{trmax} < 1.5×10^{-4} s^{-1}). This adjustment also extended the mean lifespan of detected vortices by approximately 3 hours. Under the 6-hour minimum lifespan criterion—used to filter transient disturbances—this extension nearly doubled the detection rate of moderately weak vortices (1.5×10^{-4} s^{-1} < ζ_{trmax} < 2×10^{-4} s^{-1}), highlighting the importance of ζ_{max0} in capturing less intense but persistent systems.

Second, examining the role of γ (by comparing Experiment a, which uses a lenient γ, with Experiment b, which uses an intermediate γ) revealed that the lenient γ threshold in Experiment a increased the count of weak-to-moderate vortices (1.5×10^{-4} s^{-1} < ζ_{trmax} < 3×10^{-4} s^{-1}). This increase was attributed to enhanced sensitivity to vortex splitting events, though it came with a trade-off: the mean lifespan of detected vortices was reduced by approximately 2 hours, likely due to more frequent identification of short-lived sub-vortices during splitting

Given the objective of constructing a comprehensive dataset capturing the full spectrum of PMCs, including weaker systems potentially omitted by stricter criteria, the parameter set from Experiment a was ultimately selected. This configuration yielded the highest number of vortex tracks, thereby ensuring the inclusion of marginally intense or transient PMCs and providing a more robust foundation for subsequent analysis. Validation of these results against established polar low datasets is presented in Section 4.

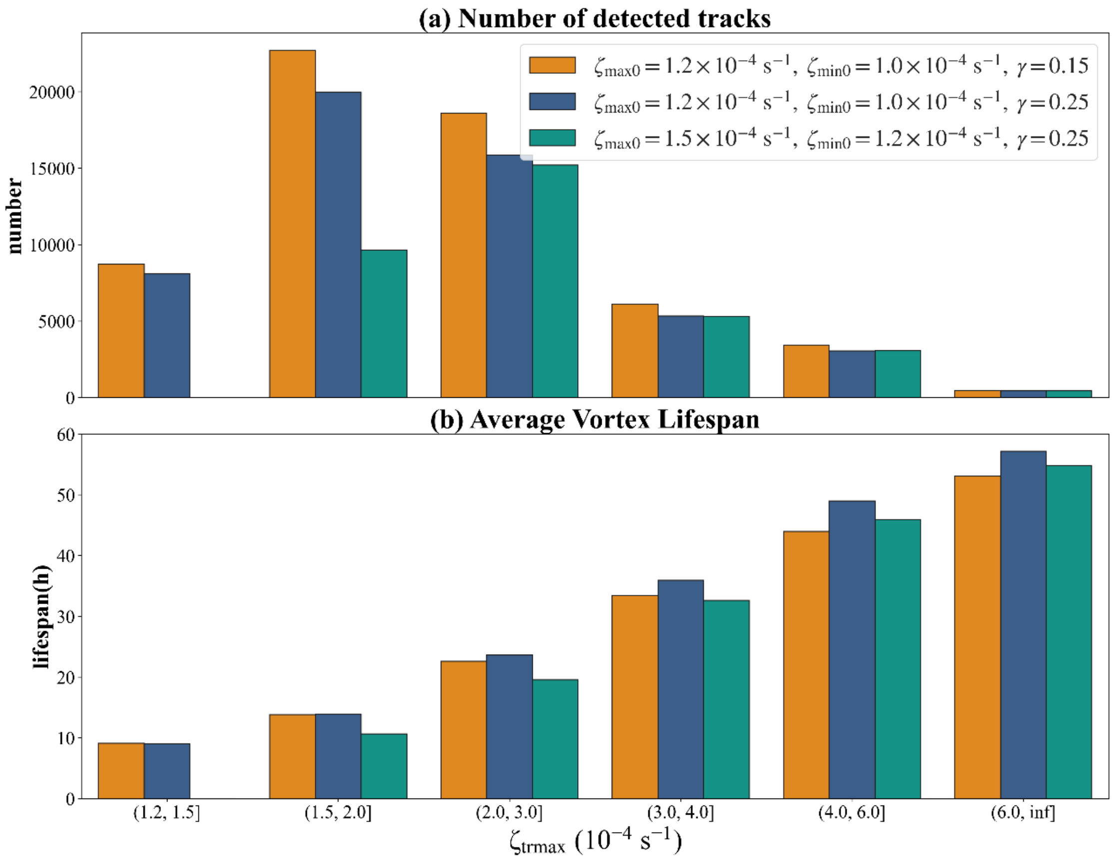

Figure 1 Sensitivity analysis of vortex identification parameters across different maximum track vorticity groups: (a) number of identified tracks, (b) mean track lifetime.

This **matching rate of the reanalysis-based track dataset with different vortex identification parameters compared to other PL track datasets** is now described in the revised version of supplement Table S2:

Experiment	Track counts	Matching rate(%) with		
		Stoll	Rojo	Noer
a	59975	93.68	69.73	87.72
b	52708	92.04	68.11	86.84
c	33622	87.39	61.35	80.70

4. It seems a YOLO (You Only Look Once) object detection algorithm is employed to detect and extract cyclonic cloud characteristics. This description of this procedure could be improved in my opinion. The authors start by generally describing the structure of the YOLOv8-obb model on line 377, with so many acronyms. However, the specific process by which this algorithm works to detect cloud features was oversimplified in the following paragraph.

Re: Thank you for your comment. We simplified the YOLOv8-obb-pose description by removing technical acronyms (e.g., decoupled head module) and retained only the framework overview. Algorithm details are deemphasized as YOLO is established.

Meanwhile, we have supplemented detailed examples of the algorithm's recognition results to help readers understand and reproduce the relevant recognition process more easily, as shown in Figure 10. **[Line 412-416]:**

The network architecture of the YOLOv8-obb-pose model comprises three main components: Backbone for multi-dimensional feature extraction, Neck for enabling multiscale feature fusion, and Head for extracting cyclone type, center coordinates, and oriented bounding box parameters (e.g., length, orientation). As shown in Fig. 10, the YOLOv8-obb-pose model successfully detects two spiral clouds (Fig. 10a) and two comma-shaped clouds (Fig. 10b) in VCI images, with oriented bounding boxes,cyclone type and center points marked.

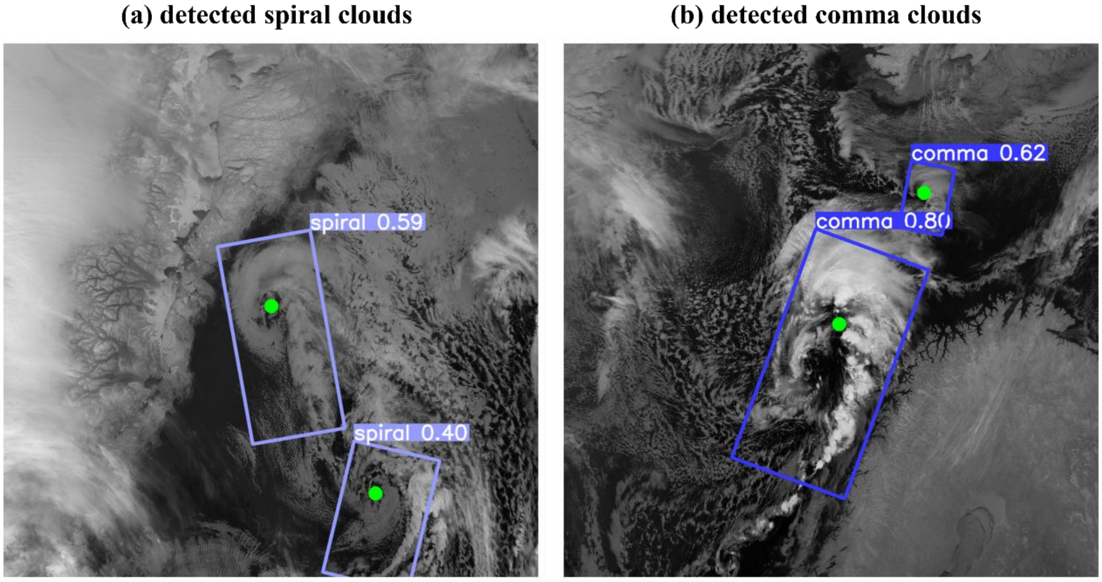

Figure 2: Examples of cyclonic cloud detection using the YOLOv8-obb-pose model: (a) two spiral clouds detected in a VCI image and (b) two comma-shaped clouds detected in a VCI image. The oriented bounding boxes for spiral clouds are shown in purple, and for comma-shaped clouds in blue. The centers of the cyclones are marked with green points. The cyclone type and detection confidence are displayed above each bounding box.

Additionally, a description of how to process the detection results to extract cyclone information is added. This helps clarify the role of the YOLOv8-obb model within the overall algorithmic workflow. **[Line 427-431]:**

For each detected cyclone, the center coordinates and the four vertices of the oriented bounding box are converted back to geodetic coordinates using the inverse of Eq. (1) and (2). The lengths of the four sides of the bounding box are calculated using the haversine formula, with the cyclone's length (width) defined as the mean size of the two long (short) sides of the rectangle. The geographic coordinates of the cyclone center are then used for subsequent matching with vortex centers.

5. When comparing the results from the IMPMCT to existing identified PL lists from previous studies, the authors give the difference in parameters and plot them. It is more appropriate to conduct a significance test between two samples in order to statistically validate the accuracy.

Re: Your opinion is very important. In comparing with other datasets, in addition to parameter difference indicators, consistency is also an important verification goal. Since there is no absolutely accurate true value dataset, we adopted the Bland-Altman analysis method for comparison. This method provides an intuitive and easy-to-understand way to evaluate the consistency of measurement values of the same object obtained by different technical means. If the distribution of differences between the two measurement results is normal, 95% of the differences should be within ±1.96 times the standard deviation of the differences, and we call this interval the 95% limits of agreement. This method evaluates the degree of agreement between the two methods by quantifying the mean difference (bias) and limits of agreement (LoA). If the vast majority of differences fall within the limits of agreement, it can be considered that the two methods have good consistency.

Results show:

- 95% of differences in vortex/cyclone properties fall within ±1.96 SD of the mean difference (Sec. 4; Table 3; Fig. S1).
- Small biases exist (e.g., mean difference: −6.8 km in vortex diameter; 0.3 hPa in SLP), attributable to methodology differences. **[Line 594-599]**.

This additional consistency test is now described in the revised version of the manuscript:

To statistically validate the agreement between IMPMCT and the Stoll (2022) dataset and Rojo list regarding vortex and cyclone properties, we performed Bland-Altman analysis (Bland and Altman, 1999). This method assesses the agreement between two measurement techniques by quantifying the mean difference (bias) and the limits of agreement (LoA), defined as the mean difference ± 1.96 standard deviations of the differences. A summary of the Bland-Altman analysis for key properties is presented in Table 3. The corresponding Bland-Altman plots, illustrating the distribution of differences against the average values for each property, are provided in Supplementary Fig. S1. As shown in Table 3, the vortex properties derived from ERA5 reanalysis data exhibit a slight systematic bias compared to other datasets. This bias is likely attributable to differences in computational methods. Critically, the Bland-Altman analysis confirms strong agreement, with approximately 95% of the differences for each property falling within the respective 95% limits of agreement (Table 3, last column), supporting the consistency between the datasets.

Table 1 Property difference between IMPMCT and other PLs list

Property	Matched number	Mean Difference	Standard Deviation of Differences	% Points within LoA
850 hPa relative vorticity (10^{-5} s^{-1})	21281	0.61	2.15	95.1
SLP (hPa)	14522	0.3	0.76	95.7
vortex equivalent diameter (km)	21281	-6.8	39.46	93.7
track-max near-surface wind speed (m s^{-1})	42	-0.27	4.83	95.2
cyclone cloud diameter (km)	892	6.76	121	94.7

This additional consistency test plotting is now described in the revised version of the supplyment:

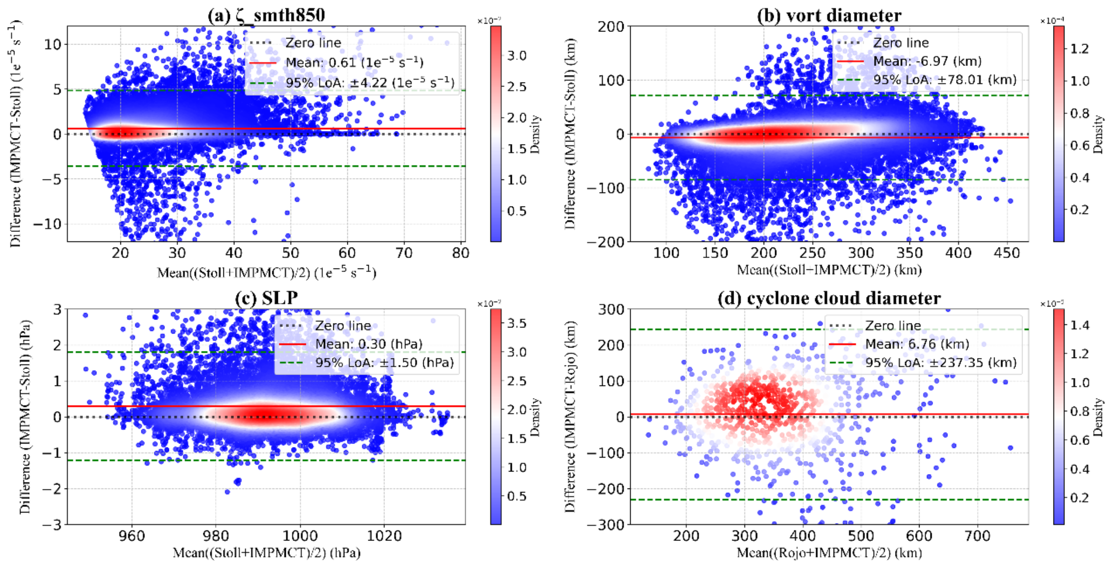

Figure S3 Bland-Altman analysis of Property Differences Between IMPMCT and Other PL list.(a) 850 hPa relative vorticity, (b) vortex equivalent diameter, (c) SLP, (d) cyclone cloud diameter.The x-axis represents the mean property value of IMPMCT and the other dataset; the y-axis represents the difference in properties (IMPMCT minus PL list). Point color indicates Gaussian kernel density. The black dashed line denotes the zero line. The red solid line indicates the mean difference of the sample properties. The upper and lower green dashed boundaries represent the limits of agreement (LoA), defined as the mean difference ± 1.96 standard deviations of the differences.*Note: Differences for properties (a), (b), and (c) are comparisons between IMPMCT and the Stoll (2022) dataset, whereas (d) uses the Rojo list. The difference analysis for track-max near-surface wind speed is not shown due to insufficient sample size.

6. Figure issues

- Specify what is plotted in Figure 1 in the name of the colorbar, same comments for Figure 3b, and Figure 7.
- The green star symbols denoting the local vorticity maxima are hard to read when overlaid on the AVHRR infrared imagery. Please change the color or enlarge the symbols. Same comments for stars in Figure 10b and wind vectors in Figure 11.
- The unit of the colobar in Figure 7a should be $1e^{-4}s^{-1}$

Re:

- Colorbar labels added to Figs. 1, 3b, 8.
- Symbols were enlarged for visibility (Figs. 1, 8, 10b, 12).
- Unit corrected in Fig. 8a to $10^{-4}s^{-1}$.

These figures have been modifiedied in the revised version of the manuscript:

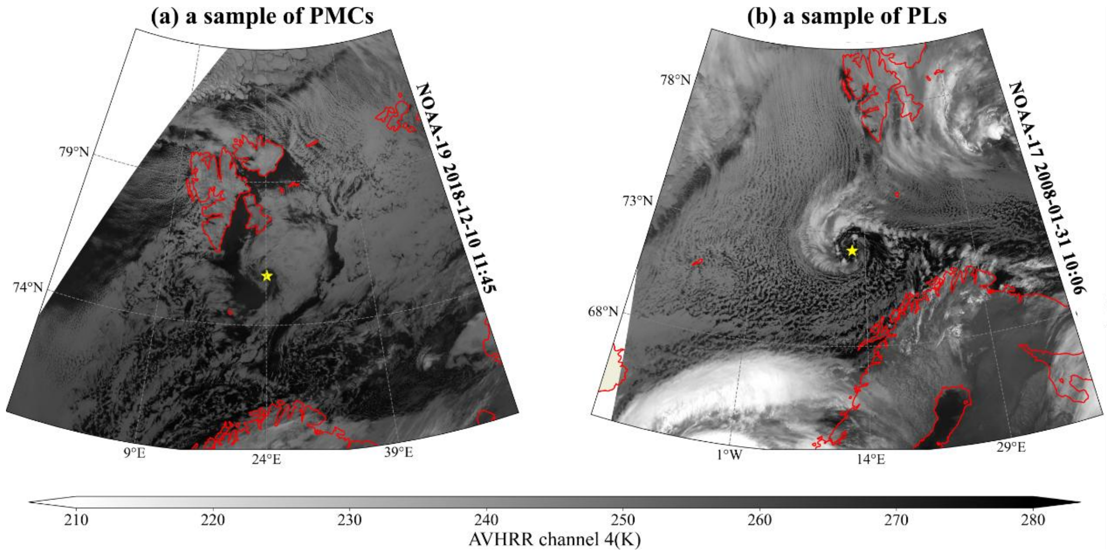

Figure 4: Two AVHRR satellite images. (a) A PMC in Barents Sea. (b) A PL in Norwegian Sea. The yellow stars mark the centers of these two cyclones.

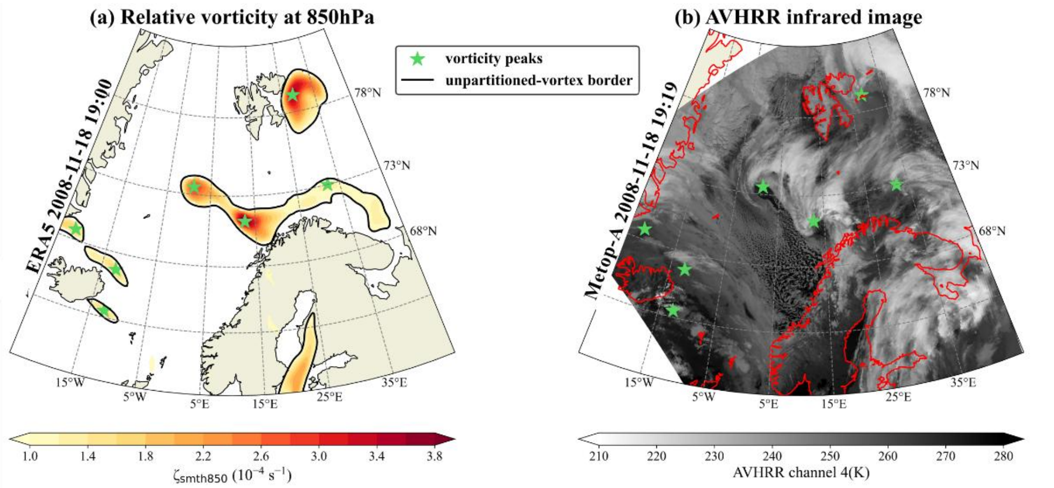

Figure 5: (a) 850 hPa relative vorticity field obtained by ERA5 data. (b) AVHRR infrared imagery concurrent with the time step in (a). The shading represents 850 hPa relative vorticity smoothed over a uniform 60 km radius and local vorticity maxima are marked by green star symbols, while regions enclosed by solid black contours denote the unpartitioned-vortex zone.

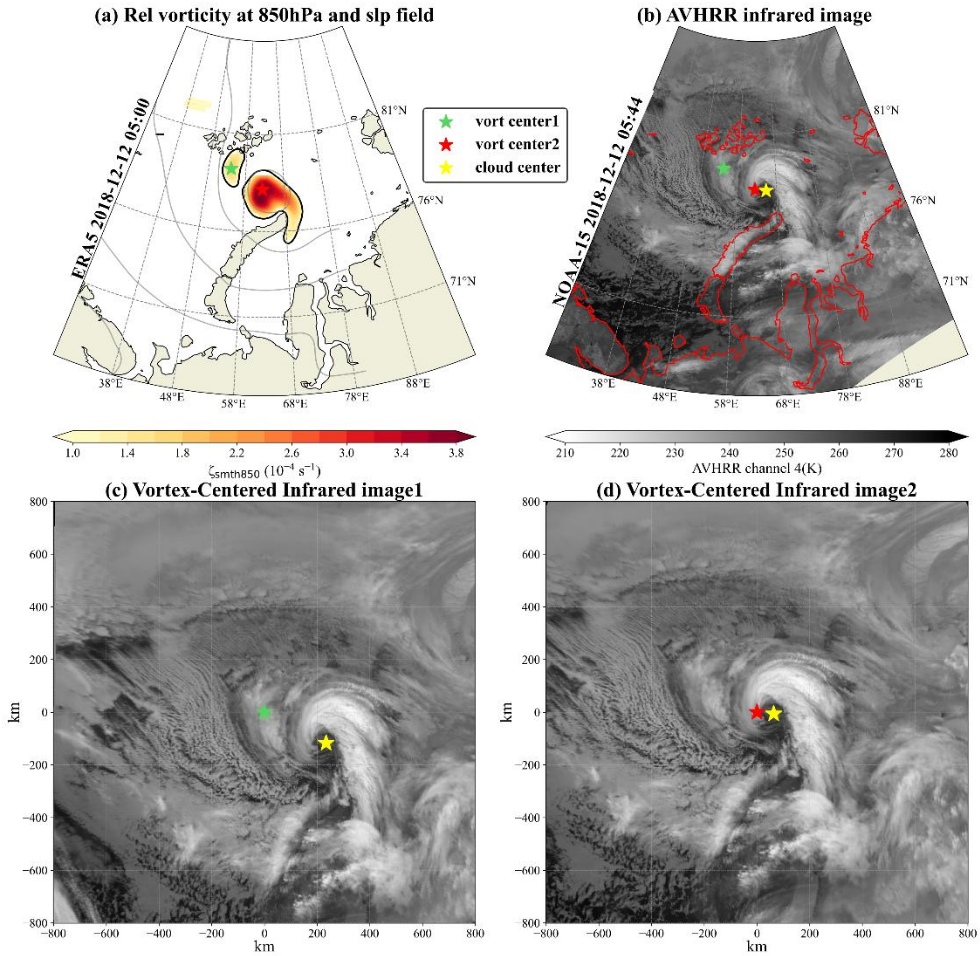

Figure 6: Two examples of VCI image generation. For the two vortices shown in (a), the AVHRR IR image (b) reveals a polar low located to the east of vortex 1 and vortex 2. This polar low exists simultaneously in the VCI images centered on vortex 1 and vortex 2 (c, d). The shading in (a) represents 850 hPa relative vorticity smoothed over a uniform 60 km radius, with gray contour lines indicating sea-level pressure at 10 hPa intervals. The centers of vortex 1, vortex 2, and the polar low are respectively marked by green, red, and yellow stars.

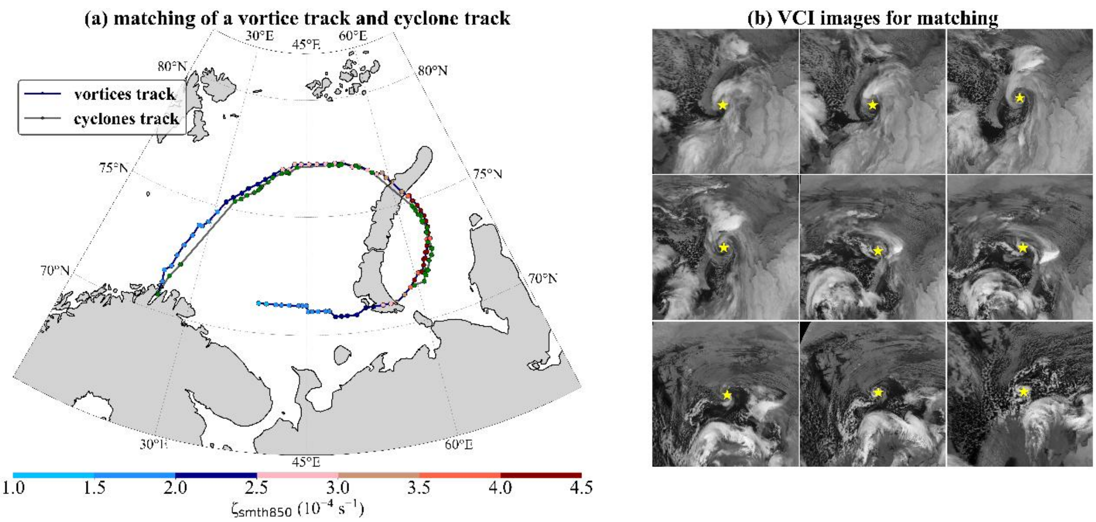

Figure 7: (a) A matched vortex track and cyclone track and (b) partial corresponding VCI images. For (a), blue solid line represents the vortex track at hourly resolution, while grey solid line with green points depicts the cyclone track points formed in VCI images that correspond one-to-one with vortex points. The color of the track points indicates the magnitude of relative

vorticity at each vortex point. For (b), the cyclone develops sequentially from left to right and top to bottom, with scan intervals between images approximately six hours apart.

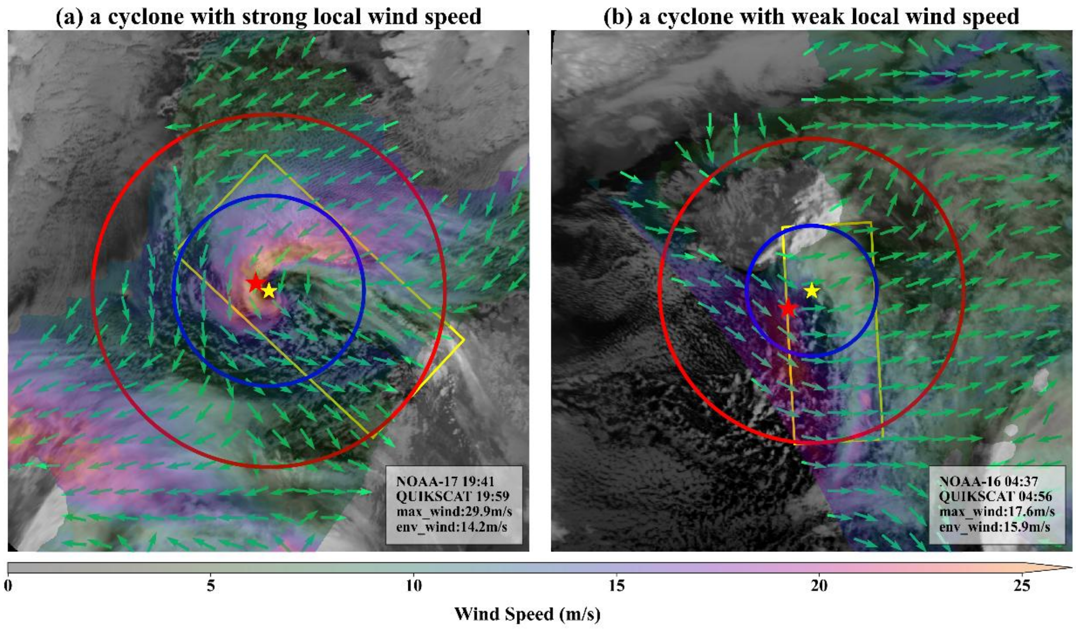

Figure 8: VCI images overlaid with near-surface wind speeds for cyclones exhibiting strong (a) and weak (b) local impacts on near-surface wind conditions. Color shading represents QuickSCAT-measured 10m near-surface wind speeds, with green arrows indicating corresponding wind vectors. Yellow borders denote the cyclones' bounding oriented box. Blue and red circular borders respectively represent the short and long search ranges. Yellow and red stars indicate the cyclone center and maximum wind speed point locations.

Minor comments:

1. Lines 41-42: Add references about this statement.

Re: We have supplemented two relevant references and revised some expressions at **[Line 41-43]**:

"Polar Mesoscale Cyclones (PMCs) are mesoscale cyclonic weather systems that frequently occur over open waters or sea-ice edges in regions poleward of the main polar front zones (Condron et al., 2006; Rasmussen and Turner, 2003)."

2. Lines 59-61: Add references about this statement or remove it as it seems irrelevant to the core points of this paragraph.

Re: We have removed the initial broad statement about the effectiveness of remote sensing. Starting directly with the core distinction criteria better aligns with the paragraph's main purpose: "Cyclonic cloud morphology and surface wind fields serve as the primary criteria..." **[Line 59-60]**.

3. Lines 129-131: Moreover, fundamental questions persist regarding the differences in formation mechanisms between PMCs and PLs, and whether PMCs can transition into PLs under specific meteorological conditions. This question seems not to be addressed.

Re: The speculative sentence on PMC-PL transition mechanisms was deleted. **[Line 127]**

4. Line 138: Winter should be defined here rather than in the Data part.

Re: The seasonal coverage of the data has been added to both the Abstract **[Line 23]** and Introduction **[Line 135].**

5. Line 140: “multi-dimensional” to “multiple”

Re: Done**. [Line 144]**

6. Line 161: “sourced” to “obtained”

Re: Done**. [Line 158].**

7. Line 169: delete “for atmospheric, land, and ocean variables”

Re: Done**. [Line 166].**

8. Lines 191- 192: Notably, QuikSCAT data spans only 1999–2009, while ASCAT has remained operational since 2010. Rephrase to: QuikSCAT operated from 1999 to 2009, whereas ASCAT has continued operations since 2010.

Re: Done**. [Line 200].**

9. Lines 281-284: “Specifically, for a vortex at a given time step, its ideal point after experiencing a time step under the steering wind influence is first calculated A search radius of 180 km is then applied around this estimated location to facilitate vortex tracking in subsequent time steps..” Should be two separate sentences.

Re: Done. **[Line 281]**

10. Lines 293-294: Rephrase to: If no spatially connectable vortices are identified in adjacent time steps, the vortex is classified as being terminated.

Re: Done. **[Line 291-292]**

Lines 316-319: Rephrase to: Building upon the lenient vorticity identification criteria established in prior analysis, a substantial population of vortex tracks has been identified within the reanalysis dataset. This collection encompasses not only cyclonic systems but also terrain-induced shear flows, low-pressure troughs, and small-scale atmospheric disturbances.

Re: Done. **[Line 349-352]**

12. Line 373: Delete “deliberately”

Re: Done. **[Line 404]**

11. Lines 391-393: Rephrase to: To ensure prediction stability, particular emphasis is placed on maintaining consistent oriented bounding box annotations and center point positions across similar evolutionary phases of cyclonic cloud morphologies.

Re: Done. **[Line 419-421]**

12. Linee 409-413: Rephrase to: To remove duplicate records, we implement a selection criterion: for any cluster of detections from the same AVHRR infrared scan (with cyclone centers <50 km apart), only the detection whose center is nearest to the VCI image center is retained.

Re: Done. **[Line 445-448]**

13. Lines 453-455: Rephrase to: To reduce the influence of strong winds in the cyclone core, we use the 75th percentile of wind speeds within the extended search radius as the environmental advection speed (reference value).

Re: Done. **[Line 488-490]**

14. Lines 484-485: Rephrase to: All reference datasets are spatially and temporally co-located with our derived tracks, retaining only those persisting for ≥3 hours.

Re: Done**. [Line 517-518]**

15. Line 526: “extraneous” to “irrelevant”

Re: Done. **[Line 563]**

16. Line 545: Rephrase to: Additionally, since the dataset includes remote sensing images of cyclones, users can easily verify the accuracy of cyclone properties and make necessary adjustments based on their specific use cases.

Re: Done. **[Line 581-582]**

17. Line 568: “these categories” to “them”

Re: Done. **[Line 614]**

References

Belmonte Rivas, M. and Stoffelen, A.: Characterizing ERA-interim and ERA5 surface wind biases using ASCAT, Ocean Sci., 15, 831–852, https://doi.org/10.5194/os-15-831-2019, 2019.

Graham, R. M., Hudson, S. R., and Maturilli, M.: Improved performance of ERA5 in arctic gateway relative to four global atmospheric reanalyses, Geophys. Res. Lett. , 46, 6138–6147, https://doi.org/10.1029/2019gl082781, 2019.

Han, Y., & Ullrich, P. A. The system for classification of low-pressure systems (SyCLoPS): An all-in-one objective framework for large-scale data sets. JGR Atmospheres, 130, e2024JD041287. https://doi.org/10.1029/2024JD041287,2025.

Moreno-Ibáñez, M., Laprise, R., and Gachon, P.: Recent advances in polar low research: current knowledge, challenges and future perspectives, Tellus: Series A, 73, 1–31, https://doi.org/10.1080/16000870.2021.1890412, 2021.

Wang, C., Graham, R. M., Wang, K., Gerland, S., and Granskog, M. A.: Comparison of ERA5 and ERA-interim near-surface air temperature, snowfall and precipitation over arctic sea ice: Effects on sea ice thermodynamics and evolution, The Cryosphere, 13, 1661–1679, https://doi.org/10.5194/tc-13-1661-2019, 2019