https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

Open Access Earth System Science Data Discussions

IMPMCT: a dataset of Integrated Multi-source Polar Meso-Cyclone Tracks

Runzhuo Fang[1,2], Jinfeng Ding[1,2], Wenjuan Gao[1,2], Xi Liang[3], Zhuoqi Chen[4], Chuanfeng Zhao[5], Haijin Dai[1,2], Lei Liu[1,2]

[1]College of Meteorology and Oceanography, National University of Defense Technology, Changsha, China
[2]Key Laboratory of High Impact Weather(special), China Meteorological Administration
[3]Key Laboratory of Marine Hazards Forecasting, National Marine Environmental Forecasting Center, Ministry of Natural Resources, Beijing, China
[4]School of Geospatial Engineering and Science, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Sun Yat-sen University, Zhuhai, China
[5]Department of Atmospheric and Oceanic Sciences, School of Physics, and China Meteorological Administration Tornado Key Laboratory, Peking University, Beijing, China

*Correspondence to: Jinfeng Ding (*dingjinfeng@nudt.edu.cn*)*

Abstract. Polar Mesoscale Cyclones (PMCs), particularly their intense subset known as Polar Lows (PLs), characterized by short lifespans of 3-36 hours and horizontal scales below 1,000 km, pose significant hazards to polar maritime activities due to extreme winds exceeding 15 m s^{-1} and wave heights surpassing 11 meters. These intense weather systems play a critical role in modulating sea-ice dynamics and ocean-atmosphere heat exchange. However, current understanding remains constrained by sparse observational records and overdependence on singular data sources (e.g., remote sensing or reanalysis). To address these gaps, this study presents the Integrated Multi-source Polar Meso-Cyclone Tracks (IMPMCT) dataset, a comprehensive 24-year (2001-2024) wintertime PMCs record for the Nordic Seas. IMPMCT combines vortices tracking algorithms from ERA5 reanalysis with deep learning-based detection of cyclonic cloud features in Advanced Very High-Resolution Radiometer (AVHRR) infrared imagery, while incorporating near-surface wind matching by Advanced Scatterometry (ASCAT) and Quick Scatterometry (QUIKSCAT) measurements. The dataset contains 1,184 vortex tracks, 16,630 cyclonic cloud features, and 4,373 wind speed records, with multi-dimensional attributes such as cloud morphology, core wind speed, and environmental advection wind speed. Validation demonstrates a 70-90 % match rate with existing PLs track datasets while providing more complete cyclone life cycle trajectories, more intuitive cloud imagery visualization, and a richer set of parameters compared to previous datasets. As the most comprehensive PMCs archive for the Nordic Seas, the IMPMCT dataset provides fundamental data for advancing our understanding of the genesis and intensification mechanisms, enables the development of enhanced monitoring and early warning systems, supports the validation and refinement of polar numerical weather prediction models, and facilitates improved risk assessment and safety protocols for maritime operations. The dataset is available at https://doi.org/10.5281/zenodo.15355602 (Fang et al., 2025).

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

1 Introduction

Polar Mesoscale Cyclones (PMCs) are mesoscale cyclonic weather systems frequently occurring in polar front zones, open waters, or sea-ice edge regions in polar areas. They are often identified on satellite imagery by comma-shaped or spiral cloud patterns. PMCs occur in all seasons but are most active in winter, with a lifespan of approximately one day and horizontal scales of less than 1,000 km (Harold et al., 1999). The most intense subset of these cyclonic systems, termed Polar Lows (PLs), are major hazardous weather phenomena in polar regions, characterized by average maximum wind speeds exceeding 15 m s^{-1} and extreme values surpassing 30 m s^{-1}. They can generate significant wave heights over 11 meters (Rojo et al., 2019), posing severe threats to human activities and maritime safety in high-latitude regions (Harrold and Browning, 1969; Orimolade et al., 2016). Additionally, PLs induce rapid sea-ice changes and intensify ocean-deep convection through dynamic and thermodynamic effects, producing complex regional climatic impacts (Clancy et al., 2022; Condron and Renfrew, 2013; Parkinson and Comiso, 2013). The Nordic Seas (Greenland Sea, Norwegian Sea, and Barents Sea) serve as a critical oceanic gateway connecting the Arctic and the Atlantic Oceans, marking a primary convergence zone for Atlantic and Arctic water masses and a key region shaping global circulation and climate (Smedsrud et al., 2022). The complex meteorological and oceanographic conditions in this area make it the most frequent PLs occurrence region (Stoll, 2022). Consequently, studying mesoscale cyclonic systems in the Nordic Seas holds significant value for Arctic maritime safety and climate change research.

Remote sensing data are effective for observing such high-impact weather systems. Cyclonic cloud morphology and surface wind fields serve as the primary criteria for distinguishing and categorizing PMCs and PLs. The former can be manually identified through visible or infrared imageries from passive radiometers (e.g., Fig. 1), while the latter can be estimated using scatterometer or microwave data. While PLs exhibit higher destructive potential and detection feasibility compared to broader PMCs, current dataset development efforts have predominantly targeted PLs, leaving PMCs relatively underrepresented in existing observational records. Blechschmidt et al. (2008) combined Advanced Very High-Resolution Radiometer (AVHRR) infrared imagery (Kalluri et al., 2021) with wind speed data derived from the Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data (HOAPS) (Andersson et al., 2010) to manually identify 90 PLs occurring in the Nordic Seas between 2004 and 2005. Noer et al. (2011) utilized AVHRR infrared imagery, Advanced Scatterometer (ASCAT), and Quick Scatterometer (QUIKSCAT) wind data to detect 121 PLs in the Nordic Seas over a decade (2000–2009). Smirnova et al. (2015) identified 637 PLs between 1995 and 2009 using Special Sensor Microwave/Imager (SSM/I) data for atmospheric total water vapor (TWV) content fields, near-surface wind speed fields, and AVHRR infrared imagery. Golubkin et al. (2021) employed Moderate Resolution Imaging Spectroradiometer (MODIS) and ASCAT data to identify PLs over the North Atlantic, compiling a catalog of 131 PLs between 2015 and 2017. In all PL lists derived from remote sensing data, the Rojo list (Rojo et al., 2015, 2019) is currently the longest temporally spanning remote sensing-derived PLs track dataset, providing tracks of 420 PLs occurring in the Nordic Seas from 1999 to 2019. It includes basic information such as cyclone location, size, cloud type,

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

Open Access Earth System Science Data Discussions

development stage, and maximum 10 m wind speed. The manually tracked datasets described above have provided valuable PLs information, contributing to ongoing research efforts. However, the unique high-latitude geography of polar regions creates significant observational challenges. Polar-orbiting satellites typically observe these regions at intervals ranging from tens of minutes to several hours, resulting in temporal gaps that make it difficult for manual tracking datasets to capture complete cyclone life cycles. Additionally, some PLs forming near sea-ice edges may exhibit distinct cyclonic cloud features exclusively during their transition over moisture-rich open waters (Bromwich, 1991), implying that remote sensing datasets could potentially miss capturing the initial developmental stages of such PLs. Consequently, while the Rojo list provides developmental pattern annotations for individual polar lows, the objectivity and quantitative reliability of these annotations remain constrained by the inherent limitations of remote sensing in achieving comprehensive characterization of PL evolution throughout their complete lifecycle. Furthermore, the occurrence of polar night, coupled with low contrast between sea-ice/snow surfaces and overlying clouds, further limits the detection capabilities of remote sensing (particularly visible-band remote sensing) methods for PLs.

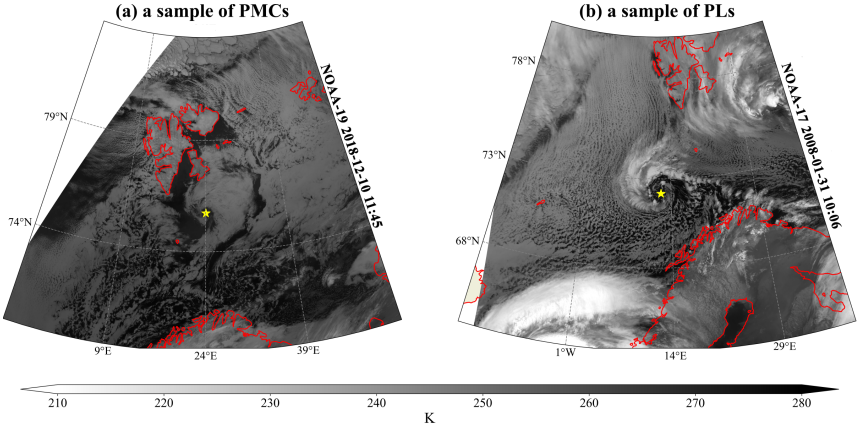


Figure 1: Two AVHRR satellite images. (a) A PMC in Barents Sea. (b) A PL in Norwegian Sea. The yellow stars mark the centers of these two cyclones.

With the improved resolution of reanalysis datasets, their ability to characterize PLs has progressively advanced (Laffineur et al., 2014; Smirnova and Golubkin, 2017), making them an increasingly critical data source for constructing PLs track datasets. Researchers have employed various combinations of identification criteria to detect PLs. For instance, Zappa et al. (2014) utilized the difference between 500 hPa temperature and near-surface temperature to represent cold air outbreak characteristics during PLs formation, while utilizing maximum near-surface wind speed to indicate PLs intensity, and 850 hPa relative vorticity to capture their cyclonic properties. Subsequent studies adopted or adapted these criteria (Stoll et al., 2018; Terpstra et al., 2016; Yanase et al., 2016). Building on the fifth-generation European Centre for Medium-Range Weather Forecasts Reanalysis (ERA5) (Hersbach et al., 2020), Stoll (2022) established a four-criteria linear-based combination defining PLs as intense mesoscale cyclones forming within polar oceanic air masses northward of the polar front. This approach successfully reproduced 60-80 % of PLs from five manual PL lists, validating ERA5's robust

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

capability in PLs representation. However, ERA5 significantly underestimates near-surface wind speeds within PL-affected regions (Belmonte Rivas and Stoffelen, 2019; Gurvich et al., 2022; Haakenstad et al., 2021), limiting its ability to objectively capture PLs' high-wind characteristics, thereby introducing notable limitations.

In summary, remote sensing and reanalysis datasets provide complementary perspectives on PLs' characteristics, with the former capturing cloud morphology and the latter resolving meteorological field distributions, highlighting their respective advantages. This complementary nature motivates the integration of both data types to construct more comprehensive PL tracking datasets—a key objective of this study. Furthermore, existing datasets primarily focus on PLs, while weaker PMCs that share similar cyclonic cloud features and environmental conditions lack comprehensive publicly available track datasets. This disparity likely stems from the fact that PMCs generally have smaller average intensities, shorter lifespans, and smaller scales compared to PLs, making them more difficult to detect. Although some researchers have proposed PMC track datasets using either remote sensing (Verezemskaya et al., 2017) or reanalysis data (Michel et al., 2018; Pezza et al., 2016; Watanabe et al., 2016), these approaches face significant limitations. Remote sensing-based datasets often have inadequate temporal coverage or lack critical near-surface wind speed records (Condron et al., 2006), while reanalysis-based datasets encounter challenges in developing effective identification criteria without remote sensing validation. As a result, no universally accepted PMC identification standards currently exist (Michel et al., 2018). Notably, while PLs have been well-documented in relation to large-scale circulation patterns such as the North Atlantic Oscillation (Claud et al., 2007) and Scandinavian blocking (Mallet et al., 2013), the climatic impacts of PMCs remain insufficiently investigated (Michel et al., 2018). 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. These knowledge gaps highlight the critical need to establish a more comprehensive tracking dataset capable of capturing PMCs throughout their lifecycle. Such a dataset would enable the complete characterization of these weaker polar mesoscale systems, representing another key motivation for this study.

Based on the above analysis, this study aims to comprehensively integrate the advantages of reanalysis datasets in characterizing the dynamical and thermodynamic structures of polar mesoscale weather systems and remote sensing data in capturing cloud morphology to establish a long-term PMCs (hereafter, "PMCs" when used alone include "PLs") track dataset in the Nordic Seas from 2001 to 2024. This dataset will contain the tracks of the PMCs in reanalysis fields and remote sensing imagery, as well as multi-dimensional attributes such as intensity, cloud morphology, and near-surface wind features. The objective is to provide a long-term, multi-attribute catalog of PMCs, offering reliable data support for atmospheric and oceanic research in the Nordic Seas.

2 Data

2.1 AVHRR data

The Advanced Very High-Resolution Radiometer (AVHRR) (Kalluri et al., 2021) is mounted on

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

NOAA series meteorological satellites and MetOp series satellites. Since its launch with the TIROS-N satellite in 1979, the sensor has continuously performed multiple daily observations of the Earth's surface. It measures reflected and emitted radiation from the Earth and its atmosphere, providing detailed information about surface characteristics, clouds, and atmospheric properties. AVHRR is an across-track scanning system with five spectral bands as shown in Table 1. It has a nadir spatial resolution of approximately 1.1 kilometers and a ±55.4° scan angle on the satellite, covering a ground swath width of 2,800 km. However, the effective resolution depends on the scan angle, with optimal image quality provided within the ±15° range.

In this study, infrared imagery used to observe cyclonic cloud features is derived from two Level 1B data products of the AVHRR (Kalluri et al., 2021): the GAC (Global Area Coverage) and LAC (Local Area Coverage) forth-band data. The GAC product provides down-sampled imagery (approximately 4 km resolution) after onboard processing, selecting every third scan line and averaging every fifth adjacent sample along the scan line. This resampling aims to ensure continuous global coverage. In contrast, the LAC product records AVHRR data at its native resolution (1.1 km) without resampling over specific orbital regions (primarily Europe and Africa), offering higher spatial resolution. All AVHRR data utilized herein are sourced from NOAA's Comprehensive Large Array-data Stewardship System (https://www.aev.class.noaa.gov/ (accessed on 18 July 2024)).

Table 1: AVHRR radiometer channel information.

Channel	Wavelength(μm)	Satellite	Application
1	0.58-0.68	ALL satellites	Surface albedo estimation
2	0.725-1.00	ALL satellites	Water body delineation
3A	1.58-1.64	NOAA15-19/MetOP A-C	Snow and ice cover identification
3B	3.55-3.93	NOAA8-19/MetOP A-C	low-level clouds identification and surface temperature
4	10.3-11.30	ALL satellites	Cloud-top temperature and surface temperature
5	11.50-12.5	NOAA8-19/MetOP A-C	Cloud-top temperature and surface temperature

2.2 ERA5 data

ERA5 is the fifth-generation global reanalysis dataset produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), designed to provide high-quality, consistent estimates of atmospheric, land, and ocean climate variables from 1950 to the present. It replaces the previous ERA-Interim dataset (Dee et al., 2011) and is currently one of the most widely used reanalysis products. ERA5 offers hourly data for atmospheric, land, and ocean variables, with a horizontal spectral truncation of T639, corresponding to a global grid resolution of approximately 31 km. The atmosphere is resolved vertically using 137 levels extending from the surface to 80 km in height.

In this study, we utilize ERA5 reanalysis data spanning 2001-2024 during the extended winter period (November-April), with a spatial resolution of 0.25° × 0.25°, covering the domain 50° N-85° N in latitude and 40° W-80° E in longitude. This dataset is employed to track vortices and compute their

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

evolutionary characteristics such as intensity and size.

2.3 QuikSCAT/ASCAT data

This study further leverages QuikSCAT and ASCAT data to examine near-surface wind field properties within the cyclone core and its surrounding ambient conditions. QuikSCAT, a NASA-developed Earth-observing satellite, employs a Ku-band SeaWinds microwave scatterometer to provide global measurements of ocean surface wind vectors. Similarly, ASCAT features a C-band microwave scatterometer aboard EUMETSAT-operated MetOp polar-orbiting meteorological satellites. These advanced instruments are specifically engineered to deliver accurate, high-resolution, continuous wind vector measurements under all weather conditions, offering comprehensive global coverage of near-surface wind patterns.

We utilize Level 2 near-surface wind vector retrieval products from both instruments to analyze wind field characteristics during cyclone development, with both datasets featuring a spatial resolution of 12.5 km. For QuikSCAT, a slice-based compositing technique integrates high-resolution measurements derived from Level 1B data into 12.5 km wind vector cells. In contrast, ASCAT employs a spatial box filter to minimize land contamination of microwave signals and enhance retrieval accuracy in coastal regions. Both datasets are sourced from NASA's Physical Oceanography DAAC (podaac.jpl.nasa.gov/ (accessed on 28 November 2024)). Notably, QuikSCAT data spans only 1999–2009, while ASCAT has remained operational since 2010. To ensure comprehensive temporal coverage across the track dataset, the two products are utilized to construct datasets across different time periods.

3 Methodology

To establish a more comprehensive cyclone track dataset in the Nordic Seas, we first utilize ERA5 reanalysis data with stable spatiotemporal resolution to obtain all vortex tracks. In this process, a lower vorticity maxima criterion is applied to extract vorticity perturbations within the reanalysis data. Subsequently, vortex tracks and their merging and splitting processes are identified based on spatial and boundary changes of vortices across consecutive time steps. For each vortex with available AVHRR data, we generate Vortex-Centered Infrared (VCI, mentioned in the following text) images to identify corresponding cyclonic cloud features with a cyclone-detection deep-learning model. Finally, near-surface wind fields derived from QuikSCAT/ASCAT are matched to characterize wind speeds within cyclone cores and environmental wind fields features. The algorithm workflow is outlined in Fig. 2, with methodological details provided in subsequent subsections.

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

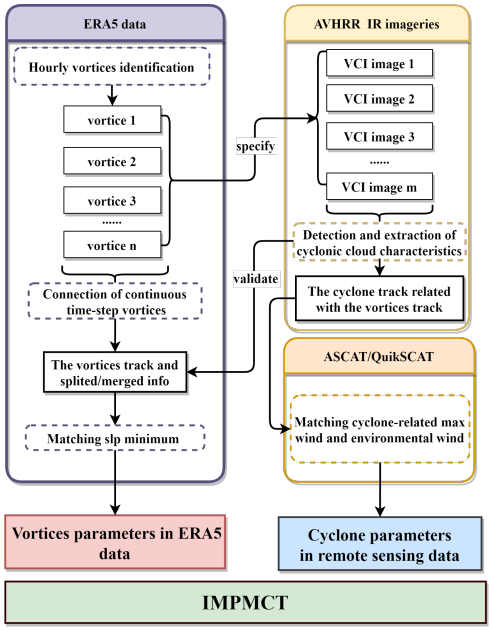


Figure 2: The workflow diagram. In the diagram, all methodologies are enclosed in dashed circular outlines, while derived datasets are framed in solid rectangular boxes. The title of each swimlane denotes the data utilized by all methods within that swimlane.

3.1 Objective algorithm for identifying and tracking vortices

Sea-level pressure (Laffineur et al., 2014; Michel et al., 2018) and low-level relative vorticity (Day et al., 2018; Stoll et al., 2021; Watanabe et al., 2016; Zappa et al., 2014) are the two most common tracer variables for PMCs in reanalysis datasets. Existing studies demonstrate that high values of low-level relative vorticity, compared to sea-level lows which are susceptible to synoptic scale pressure fields, are more closely associated with actual cyclone positions and exhibit smaller biases in cyclone detection and intensity estimation (Stoll, 2022; Stoll et al., 2020; Zappa et al., 2014). Therefore, we apply an objective mesoscale vortices-tracking algorithm to the 850 hPa relative vorticity fields in ERA5 data to obtain hourly-resolution vortex tracks. This algorithm was first proposed by Shimizu and Uyeda (2012) to track convective cells prone to merging and splitting, and has since been developed and improved for PMC tracking (Watanabe et al., 2016; Stoll et al., 2021). It specifically comprises two components: hourly vortices identification and connection of continuous time step vortices.

3.1.1 Hourly vortices identification

When multiple vortices coexist within the same region of cyclonic shear flow, they often manifest as a contiguous positive vorticity zone in the vorticity field (hereafter referred to as an unpartitioned-vortex in the algorithm). The major challenge in vortex identification within vorticity fields is how to partition such regions (as exemplified in Fig. 3) into distinct isolated vortex regions.

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

Open Access Earth System Science Data Discussions

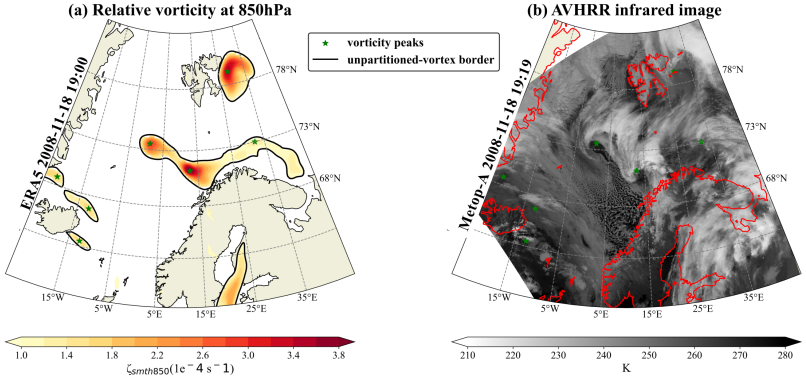


Figure 3: (a) 850 hPa relative vorticity field obtained by ERA5 data. (b) AVHRR infrared imagery concurrent with the time step in (a). The shading represents 850 hPa relative vorticity smoothed over a uniform 60 km radius and local vorticity maxima are marked by green star symbols, while regions enclosed by solid black contours denote the unpartitioned-vortex zone.

First, a uniform smoothing of 60 km is applied to the hourly 850 hPa relative vorticity field to disconnect weak vorticity continuity zones and eliminate minor perturbation maxima. Subsequently, in the smoothed vorticity field, regions enclosed by closed contour lines exceeding a minimum threshold ζ_{min0} are identified as unpartitioned vortices. Thereafter, each unpartitioned-vortex (e.g., the area within the thick black solid line in Fig. 4) is subjected to isolated vortex extraction via the following procedure:

Step 1: Identify local vorticity maxima exceeding the threshold ζ_{max0}, designated as vortex peaks with relative vorticity values ζ_{max} (e.g., in Fig. 4, three local vorticity maxima satisfy $b > a > c$). Contour lines (gray thin solid lines) are then drawn at 10^{-6} s^{-1} intervals. Subsequently, the outermost contour line enclosing each individual or combined peak(s) is identified as the valley-line (black thin solid lines, e.g., $\zeta_{min1} \approx \zeta_{min2} < \zeta_{min3} \approx \zeta_{min4}$ in Fig. 4). These valley-lines enable the separation of distinct vortex regions containing single or multiple peaks.

Step2: The isolation status of each vortex region is determined by assessing the relative disparity between each valley-line and its internal maximum peak. As illustrated in Fig. 4: peak a represents the strongest peak within its associated valley-line ζ_{min4}, peak b corresponds to the maximum within two valley-line-enclosed areas ζ_{min1} and ζ_{min3}, and peak c is the dominant peak within its respective valley-line ζ_{min2}. The assessment proceeds systematically through vortex regions in descending order of their valley-line vorticity magnitude (ζ_{min}): for the maximum peak with relative vorticity value ζ_{max} within the valley-line-enclosed vortex region, if the criterion $(\zeta_{max} - \zeta_{min})/\zeta_{max} > \gamma$ is satisfied (where γ denotes the isolation vortex threshold), the area centered on this peak and bounded by the valley-line is classified as an isolated vortex region. If a vortex region contains only one such isolated vortex region, the isolated vortex will be expanded to encompass the entire domain. (in Fig. 4, the vortex region enclosed by ζ_{min4} associated with peak a fails to meet the isolation criterion. Conversely, peaks b and c forming two distinct isolated vortex regions bounded by their respective valley-lines ζ_{min1} and ζ_{min2}).

Step3: For all vortex points located within each unpartitioned-vortex but outside the isolated vortex

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

regions, each point is assigned to the nearest isolated vortex based on geographical distance. Finally, all isolated vortices in the each unpartitioned-vortex region are mutually designated as adjacent vortices (e.g., vortices b and c), serving as inputs for subsequent analysis of merging or splitting events. The area of each vortex is defined by its corresponding allocated isolated vortex region.

To maximize the inclusion of potential PMCs, we implement more lenient vortex detection criteria compared to Stoll et al. (2021). The key modifications include: (1) reducing the vorticity peak threshold ζ_{max0} from 1.5 to 1.2 (10^{-4} s^{-1}), (2) lowering the vortex boundary threshold ζ_{min0} from 1.2 to 1.0 ($\times10^{-4}$ s^{-1}), and (3) decreasing the isolated vortex threshold γ from 0.25 to 0.15. These relaxed criteria enable the identification of weaker-intensity vortices and further extend the lifecycles of vortex tracks. For splitting or adjacent vortices, the smaller γ value makes the precursors of splitting more discernible and reduces the likelihood of erroneous merging during adjacency processes.

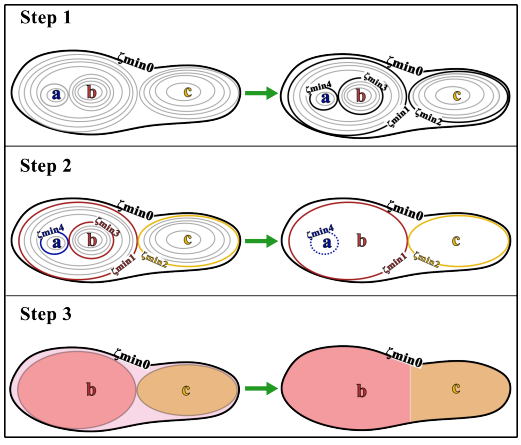


Figure 4: Vortex identification algorithm example. The black thick solid lines ζ_{min0} represent the unpartitioned-vortex border. The vorticity peaks a, b, and c are three detected vorticity local maxima within this unpartitioned-vortex. The thin black solid lines from ζ_{min1} to ζ_{min4} in Step 1 denote vortex valley-lines that divide single or multiple peak regions. After vortex isolation assessment in Step 2, the retained valley lines ζ_{min1} and ζ_{min2} for peaks b and c form the initial boundaries of their respective isolated vortices, while vortex a is classified as non-isolated, with its boundary shown as a dashed line. In Step 3, the pale pink regions outside the isolated vortices are further allocated to vortices b and c.

3.1.2 Connection of continuous time step vortices

Based on the results of hourly vortices identification, the introduction of steering wind is employed to estimate the movement of vortices. The steering wind is computed by averaging wind fields within a 450 km radius around the vortex center at 550 hPa, 700 hPa, and 850 hPa, which is statistically proven to have minimal bias (Yan et al., 2023). 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.. Subsequently, the (a) nearest neighbor principle or (b) maximum area overlap principle (as shown in Fig. 5) is applied to connect vortices between two consecutive time steps, when vortices exist within the estimated region, the nearest vortex is connected; otherwise, the vortex with the largest area overlap within the region is selected for connection. Finally, if the distance between the

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

Open Access Earth System Science Data Discussions

centers of vortices to be connected in adjacent time steps exceeds 200 km and the vorticity of the vortex center at next time step is less than 1.5×10^{-4} s^{-1}, the connection is terminated to minimize spurious connections.

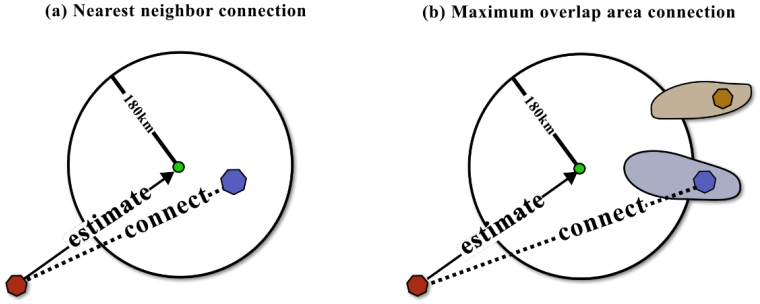


Figure 5: Schematics of continuous time step vortices connection

Additionally, if no connectable vortices are detected in adjacent time steps, the vortex is deemed to cease activity at that time step. Under the assumption of constant centroid positions during splitting and merging (Shimizu and Uyeda, 2012), if a vortex is contiguous to other vortices at its start (end) track point, it is considered to have been generated (terminated) via splitting (merging). As shown in Fig. 6, in two simplified vortex motion scenarios, vortex *b* begins splitting and merging at the t3 time step.

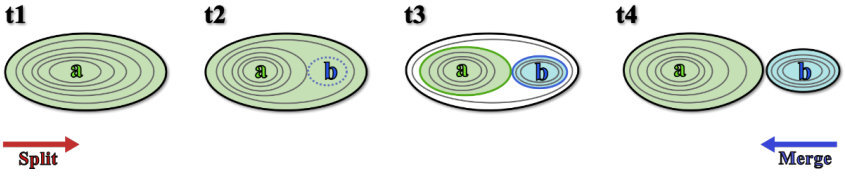


Figure 6: The schematic diagram illustrates two vortices splitting and merging processes. The t1 to t4 represent four consecutive time steps. The red/ blue arrow indicates the direction corresponding to the splitting/ merging process of two vortices. The colored regions and solid lines represent isolated vortex regions and their boundaries. Gray solid lines show contour lines of the 850 hPa relative vorticity field, and black solid lines indicate the unpartitioned-vortex boundaries. The blue dashed line indicates that the vortex *b* is not yet an isolated vortex at time t2.

3.2 Matching SLP minimum

While vortices often fail to produce closed isobars in SLP fields due to interference from background pressure gradients, their atmospheric influence can still be quantified through detectable SLP minima. Notably, certain polar lows originate within upper-level cold-core systems (known as "cold low types") frequently generate deep convection and produce substantial near-surface impacts (Rasmussen, 1981; Businger and Reed, 1989). To systematically capture these characteristics, the SLP field is first smoothed using Gaussian filtering with a radius of 50 km to suppress noise. Subsequently, the SLP minimum point located within a 150 km radius of the nearest vortex centroid is designated as the SLP center for that vortex.

3.3 Detection and extraction of cyclonic cloud characteristics

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

Building upon the lenient vorticity identification criteria previously established, a substantial population of vortex tracks have been identified using reanalysis data, including not only cyclonic systems but also terrain-induced shear flows, low-pressure troughs, and small-scale atmospheric disturbances. To validate whether these vortices belong to the PMCs, AVHRR infrared imageries are further employed for verification. This process begins with temporal matching of satellite overpasses to vortex track timesteps, followed by generation of Vortex-Centered Infrared (VCI) images through linear interpolation of infrared data onto a geographically-referenced 801×801 grid coordinate with 2 km resolution, centered on each vortex center (Fig. 7c and Fig. 7d). The coordinate transformation employs the formulas:

$$lat(x,y) = \frac{y}{2\pi R} + vort_{lat}\,, x, y \in \{-800, -798, \ldots, 798, 800\} \quad (1)$$

$$lon(x,y) = \frac{x}{2\pi R \cdot cos\,(vort_{lat})} + vort_{lon}, x, y \in \{-800, -798, \ldots, 798, 800\} \quad (2)$$

The coordinate transformation utilizes $vort_{lon}$ and $vort_{lat}$ as the longitude and latitude of the original coordinate grid, corresponding to either the vortex center at the given timestep or the VCI image center. This approach implements an equirectangular projection that provides a first-order approximation of geographic coordinates within the vicinity of the origin point.

The VCI images enable comprehensive analysis of cloud features within a 1600 km×1600 km domain centered on each tracked vortex position, providing an optimal spatial scale that captures the majority of PMCs while simultaneously accommodating larger-scale extratropical systems advected into Arctic regions. By transitioning from broad-scale satellite observations to these precisely localized domains, this imagery method significantly enhances the spatial correspondence between vorticity-derived tracks and cloud features, with particular sensitivity improvement for smaller-scale and shallower cyclones. Meanwhile, the georeferenced framework of VCI images provides two critical analytical capabilities: first, it enables direct quantification of cyclone dimensions through the standardized geographic grid; second, it allows precise measurement of positional discrepancies between observed cloud systems and modeled vortices through center-to-center displacement vectors. Furthermore, VCI images are also generated for two-time steps before the start and after the end of each vortex track. This allows us to capture the initial formation and dissipation stages of PMCs that are not adequately represented in vorticity fields, enabling users to better evaluate the representation of PMCs.

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

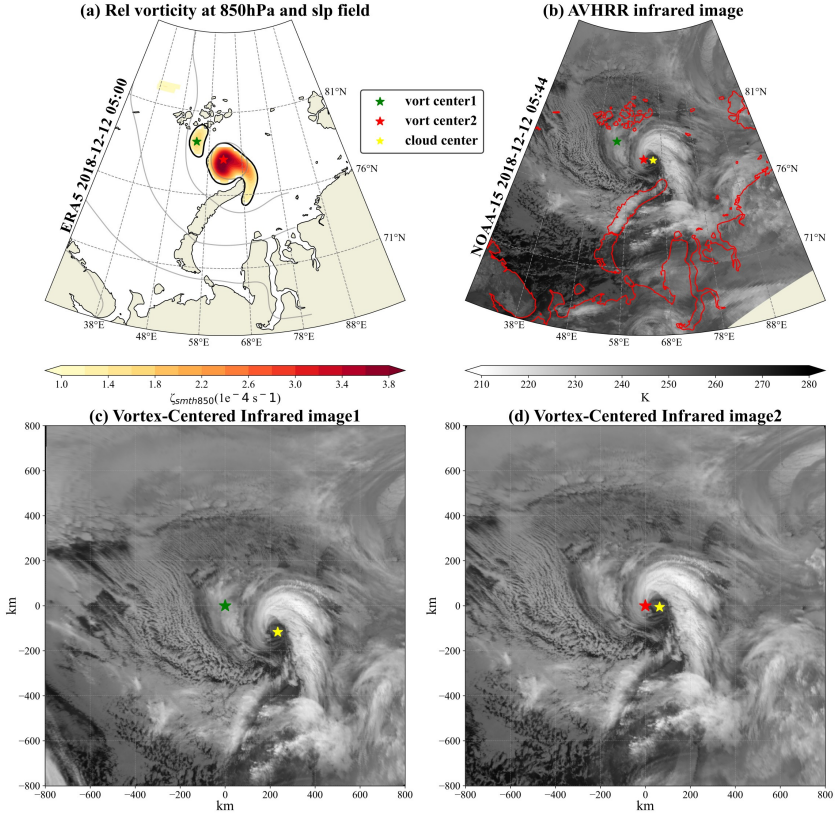


Figure 7: Two examples of VCI image generation. For the two vortices shown in (a), the AVHRR IR image (b) reveals a polar low located to the east of vortex 1 and vortex 2. This polar low exists simultaneously in the VCI images centered on vortex 1 and vortex 2 (c, d). The shading in (a) represents 850 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.

To further extract cyclonic cloud features corresponding to vortices from the vast collection of VCI images, the YOLO (You Only Look Once) object detection algorithm is employed to automate this process. Object detection is a computer vision task that uses neural networks to locate and classify objects within images. The YOLO series of algorithms (Redmon et al., 2016), characterized by high efficiency and accuracy, has become prominent in real-time object detection tasks across various fields, from agriculture to healthcare. In this track dataset construction, the YOLOv8 framework is adopted to automatically extract cyclonic cloud morphology features, including cloud type classification (spiral cloud or comma-shaped cloud), center coordinates, and an oriented bounding box enclosing the cyclone.

Figure 8 illustrates typical cyclonic cloud morphologies, the most common comma-shaped cloud structure is shown in Fig. 8a, where the head is typically composed of a tall and smooth cirrus shield surrounding a dark, nearly cloud-free center. Ripple-like wave patterns sometimes appear at the edge of the head, indicating significant wind shear within the cyclone. Fig. 8d presents the typical spiral cloud

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

morphology, characterized by one or more convective cloud spiral bands encircling the circulation center. These spiral bands are occasionally predominantly composed of cellular clouds. Intermediate baroclinic forms illustrated in Fig. 8b and Fig. 8c represent transitional stages between comma and spiral types, sharing structural similarities with occluded extratropical cyclones but at reduced horizontal scales, and are consequently classified within the spiral category. Additionally, the analytical framework of oriented bounding box is also introduced that provide quantitative measures of cyclone scale, with the long axis aligned parallel to the tail cloud band and the short axis tangent to the cloud head. While conventional approaches estimate cyclone size using the mean axis length (Smirnova et al., 2015), this dataset deliberately provides separate measurements of both axes to account for potential overestimation caused by the connection of tail cloud band of cyclones and long cloud bands of mesoscale-front, thereby enabling researchers to make more precise assessments of true cloud coverage dimensions.

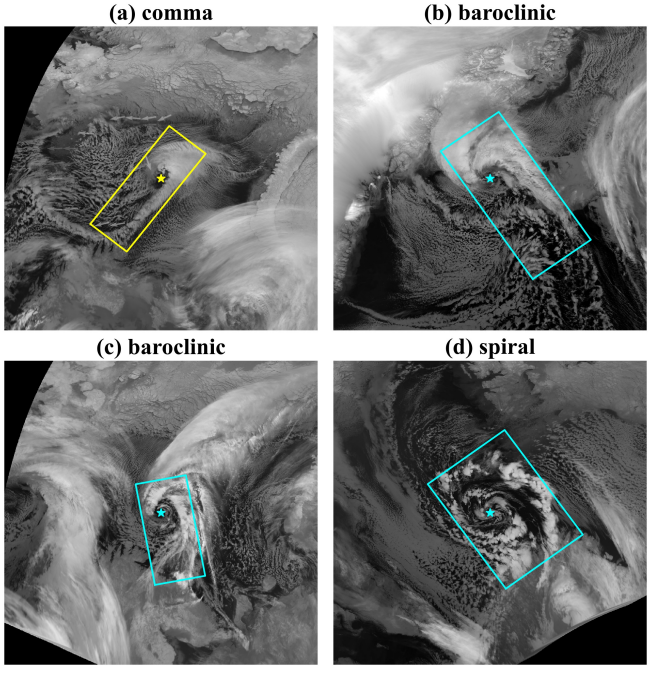


Figure 8: Different cyclonic cloud morphologies in four VCI images: (a) comma-shaped cloud; (b), (c) and (d) spiral clouds. The yellow/blue bounding boxes and stars respectively denote the oriented bounding boxes and center positions of comma-shaped/spiral cyclones.

To identify the aforementioned cyclonic cloud features, the YOLOv8-obb-pose model is configured using the YOLOv8 model framework (Jocher et al., 2023), which combines oriented object detection (obb) and keypoint detection (pose). Specifically, a branch for keypoint prediction is added to the decoupled head module of the YOLOv8-obb model. This enables the new YOLOv8-obb-pose model to simultaneously perform automatic detection of cyclone type, center position, and oriented bounding box. The network architecture of the YOLOv8-obb-pose model, as shown in Fig. 9, comprises three main components: Backbone, Neck, and Head. The Backbone uses the CSPDarknet

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

structure, which includes Conv Block (basic convolutional units for feature extraction), C2F Block (enhancing multi-scale and multi-dimensional feature extraction), and SPPF (capturing hierarchical contextual information). The Neck employs the PAFPN (track Aggregation Feature Pyramid Network) to construct a feature pyramid of VCI images, enabling multiscale feature fusion. The Head adopts a decoupled architecture with parallel branches for extracting object category (cyclone type), center coordinates, and oriented bounding box parameters (e.g., length, orientation).

During the model training process, we first construct a manually annotated dataset to train the YOLOv8-obb-pose model. Particular attention is given to maintaining consistent oriented bounding box annotations and center point positions across similar evolutionary phases of cyclonic cloud morphologies to ensure prediction stability. To optimize the trade-off between detection efficiency and accuracy, we implement an iterative training protocol involving successive cycles of prediction, manual correction, and retraining using VCI images. As detailed in Table S1, the model achieves competitive performance metrics on the validation set following this optimization process. The final YOLOv8-obb-pose implementation demonstrates robust capabilities in both cyclone detection and center localization tasks, satisfying requirements for practical applications.

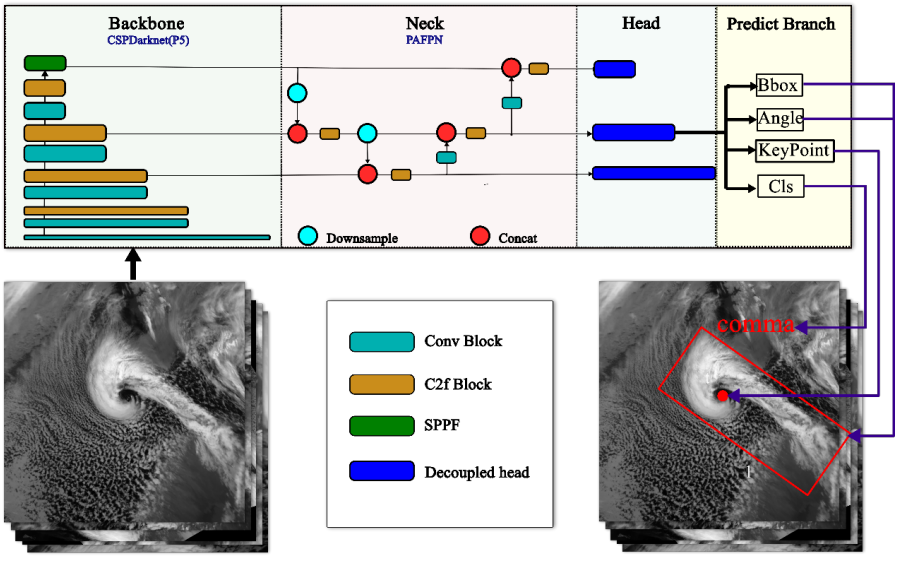


Figure 9: YOLOv8-obb-pose model schematics.

3.4 Validation of the vortex tracks

Each series of VCI images based on vortex track provides spatiotemporal neighboring local infrared cloud imagery that follows the vortex's movement. After extracting cyclonic features from VCI images, whether a vortex track corresponds to a cyclone evolution process is determined by proximity matching between the cyclone center detected in each VCI image and the vortex center. The following steps ensure that each VCI image only retains a cyclone uniquely matched to a vortex track point:

I Uniqueness: As illustrated in Fig. 7, spatially proximate vortices in reanalysis data can result in

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

multiple detections of the same cyclone across corresponding VCI images. To eliminate duplicate records, we establish a selection criterion: for any cluster of detections originating from the same AVHRR infrared scan (where cyclone centers are separated by less than 50 km), only the cyclone whose center is closest to the VCI image center is retained.

II Proximity: Each VCI image retains only the cyclone whose center is nearest to the VCI image center and within 250 km of it.

Further, we extend the point-to-point matching to the track-to-track. When points of a vortex track are continuously matched with a series of cyclones in VCI images over four hour or longer time-steps with an average matching distance less than 150 km, the vortex track is preliminarily identified as a PMC track associated with the cyclone evolution process (as shown in Fig. 10)

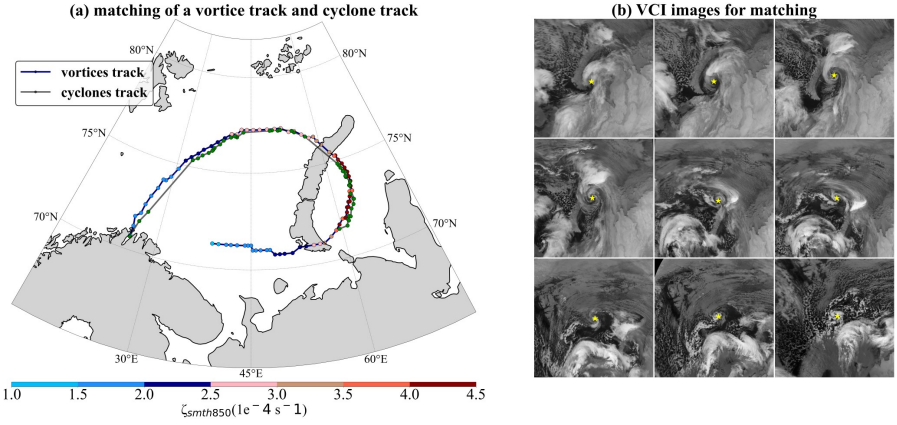


Figure 10: (a) A matched vortex track and cyclone track and (b) partial corresponding VCI images. For (a), blue solid line represents the vortex track at hourly resolution, while grey solid line with green points depicts the cyclone track points formed in VCI images that correspond one-to-one with vortex 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.

3.5 Matching cyclone-related max wind and environmental near near-surface wind

When cyclonic cloud features are detected in VCI images, near-surface wind speeds over the ocean are further matched to estimate their intensity. According to existing definitions (Rasmussen and Turner, 2003), PLs typically produce high near-surface wind speeds exceeding 15 m s^{-1} (gale force), with high-wind-speed regions concentrated in narrow cloud bands linked to the eye wall or intense convective zones encircling the center. In contrast, relatively weaker PMCs generally fail to penetrate the temperature inversion layer above the marine mixed layer, resulting in smaller near-surface wind speeds (Noer et al., 2011).

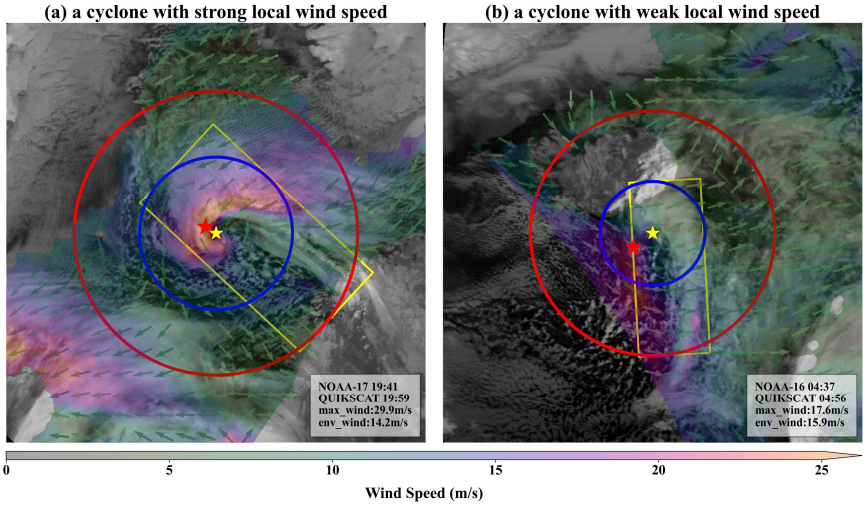
However, while environmental advection speed is not mentioned in existing definitions, some weaker PMCs forming during cold air outbreaks often exhibit wind speed maxima exceeding 15 m s^{-1} due to the cold air advection speeds, which are frequently subjectively selected to avoid misclassification as PLs (Wilhelmsen, 1985). This underscores the importance of environmental advection speed as a key environmental factor in studying such systems. In this study, we incorporate environmental advection speed into the track dataset using a cloud-scale-based search radius method to

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

facilitate user reference.

For each VCI image, ASCAT/QuikSCAT wind speed data is selected for matching to generate the cyclone core maximum near-surface wind speed and environmental advection wind speed when the time difference between the wind data and VCI image scan is less than 30 minutes. Given that the majority of PLs have movement speeds below 13 m s^{-1} (Rojo et al., 2015; Smirnova et al., 2015), the representative error caused by this 30-minute time window is considered negligible. For the matched wind speed data, we establish two wind speed search radii surrounding the cyclone center, which are significantly correlated with the cyclone cloud system scale, to distinguish cyclone core near-surface wind speeds and environmental advection wind speeds. The short search radius is defined as the distance from the cyclone center to the nearest short edge of its bounding oriented box, confining the wind speed search area to the high-wind-speed region near the cyclone's head, with the maximum wind speed within this area designated as the cyclone's max wind speed. The long search radius is set as the distance from the cyclone center to the farthest short edge of the bounding oriented box, used to sample the environmental wind field surrounding the cyclone. To minimize influence from the cyclone core's strong winds, the 75th percentile of all wind speed data within the long search radius is adopted as the environmental advection speed (reference value).

For data users, environmental advection speed provides additional characteristics of cyclone intensity. When the difference between the cyclone's maximum wind speed and the environmental advection speed is positive and large, it indicates the cyclone possesses a higher destructive potential, whereas a negative or smaller difference suggests the cyclone's maximum wind speed is dominated by advection speed. The former scenario is exemplified in Fig. 11a, where a strong wind speed zone far exceeding the cold air advection speed is evident at the cyclone's head, highlighting the cyclone's significant impact on near-surface wind conditions. Conversely, Fig. 11b demonstrates a cyclone lacking a well-developed cyclonic circulation, instead exhibiting wind characteristics consistent with the environmental wind field.



https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

Figure 11: 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.

4 Results and discussion

Our analysis begins by applying a vortex tracking algorithm to reanalysis data, which identifies 57,688 vortex tracks. Through validation with VCI images, we confirm 1,184 cyclone-related vortex tracks containing 16,630 distinct cyclone cloud features. Subsequent analysis of surface wind speed characteristics reveals 4,373 instances with measurable wind patterns, including 845 tracks where maximum wind speeds exceeded the 15 m s^{-1} threshold. These validated 1,184 vortex tracks and their associated remote-sensing images constitute the IMPMCT track dataset. The accuracy of IMPMCT is rigorously evaluated through comprehensive comparisons with existing manually identified and reanalysis-based track datasets.

First, to validate the accuracy of the vortex track datasets obtained from the vortices tracking algorithm, they are compared with the manually identified PL lists published by Noer et al. (2011), Rojo et al. (2019), and the objectively derived PL track datasets from reanalysis data by Stoll (2022). All reference datasets are spatially and temporally constrained to match our tracks, with only tracks persisting for more than 3 hours retained. We implement the following matching criteria: a vortex track is considered matched with a PL track when more than 50 % of temporally coincident track points (within ±1 hour) are located within a 150 km radius (applying an 80 % threshold for Stoll's dataset). A vortex track is allowed to match multiple PL tracks from reference datasets, provided that these PL tracks have no overlapping segments and each PL track is uniquely paired with its nearest vortex track. As presented in Table 2, the validation results demonstrate strong agreement with Stoll's dataset, confirming the robustness of our vortex tracking algorithm; not only that, we achieve higher matching rates with manual PL lists by using lower vortex identification thresholds, which to some extent further demonstrates the enhanced representation capability of ERA5 reanalysis data for PLs characterization.

Table 2: the matching rate of the reanalysis-based track dataset for IMPMCT generation compared to other PL track datasets.

PL tracks	Time period	Tracks in Nordic Sea with life exceeding three hours	ERA5 matched	Matched fraction %
Noer	2002-2011	114	100	87.72
Rojo	2000-2019	370	259	70.00
Stoll	2000-2020	3179	2881	90.63

After ensuring the validity of the reanalysis-based vortex track dataset for IMPMCT generation, we further analyze the reliability of vortex properties in IMPMCT by comparing three characteristic parameters (850 hPa relative vorticity, SLP center, and vortex equivalent diameter) with those provided

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

Open Access Earth System Science Data Discussions

by Stoll's dataset, along with the distance between vortex centers across the two datasets. IMPMCT retained 678 matched tracks with Stoll's dataset. As shown in Fig. 12a, among the matched tracks, 90 % of vortex points remain within 50 km of each other at the same time step. Additionally, the average absolute differences of the three vortex properties at these proximate track points remain small: 1.09×10^{-5} s^{-1} for relative vorticity, 0.43 hPa for sea-level pressure, and 23.21 km for vortex equivalent diameter. Additionally, vortex property differences increase with distance, indicating that discrepancies between IMPMCT and Stoll tracks stem from differing identification thresholds. To further demonstrate that such discrepancies are not errors, we calculate the standard deviation of vortex properties across three consecutive time steps for each track and computed track-wide averages. If the calculated property standard deviations remain low, it indicates stable variations in vortex properties along the track, suggesting reliable track identification. Fig. 12 (b), (c) and (d) show the track-averaged local standard deviations of the three vortex properties for IMPMCT and Stoll datasets. The local variabilities in IMPMCT tracks generally align with Stoll's variabilities and are even smoother in some cases.

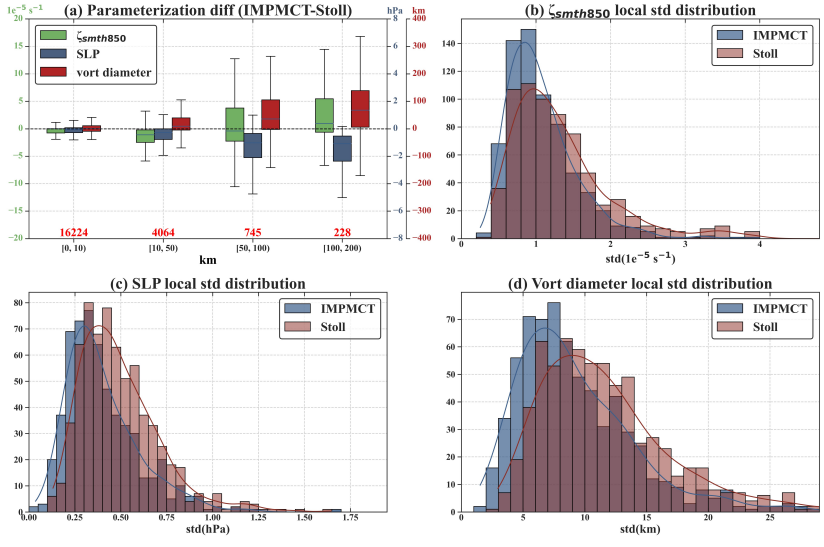


Figure 12: Distribution of differences in three vortex properties and their track-averaged local standard deviations at co-located hourly track points between matched IMPMCT and Stoll tracks. The boxplot in (a) shows property differences as a function of spatial deviation distance between matched track points. The red numbers above the x-axis indicate the count of track point pairs in each distance bin. Each boxplot's y-axis scale corresponds to the color of its respective property (green: relative vorticity, blue: sea-level pressure, red: vortex diameter). Frequency histograms and fitted curves of track-averaged local standard deviations for the three properties are displayed in (b) relative vorticity, (c) sea-level pressure, and (d) vortex diameter.

IMPMCT uses hourly-resolution vortex tracks from reanalysis data as a supplement to cyclone tracks. The correspondence between vortex and cyclone tracks is established solely through continuous spatiotemporal matching of their respective centers. To ensure the accuracy of this correspondence, we perform subjective validation to confirm that each cyclone track does not incorporate extraneous cyclonic processes. Notably, while the average matching distance between vortex and cyclone tracks is

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

constrained within 150 km, approximately 95 % of track pairs have average matching distances below 100 km (as shown in Fig. 13), demonstrating strong consistency between cyclone and vortex tracks.

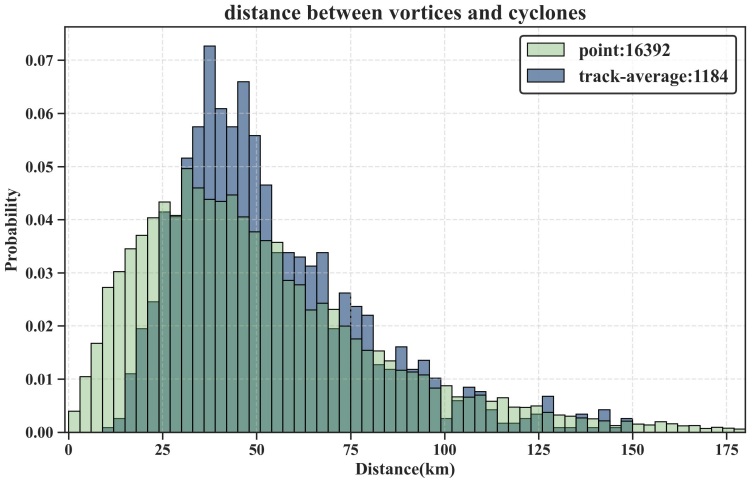


Figure 13: Probability distribution of distances between matched cyclone-vortex points (green) and track-average distances (blue).

The cyclone properties in IMPMCT primarily include cyclone scales and maximum core near-surface wind speeds. These properties are evaluated through comparison with the Rojo list. For cyclone scale validation, we compare the diameter from the Rojo list with the approximate cyclone scale in the IMPMCT dataset (calculated as the average of cyclone width and length). We match cyclone tracks between IMPMCT and Rojo list based on the following criteria: the nearest cyclone centers are matched if their distance is less than 120 km and their central scan times are within 60 minutes. A cyclone track pair is deemed matched if the number of matched points exceeds half the total points of a PL track in the Rojo list. Comparisons of cyclone cloud scale and maximum wind speeds between the matched time periods are shown in Fig. 14. Despite fewer matched tracks due to differences in data sources, the cyclone size and wind speed attributes between IMPMCT and Rojo lists remain largely consistent. Considering the differences in remote sensing data and subjective criteria, we consider such discrepancies reasonable. Additionally, since remote sensing images of cyclones are provided in the dataset, users can conveniently verify the accuracy of cyclone properties and apply appropriate corrections based on their specific use cases.

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

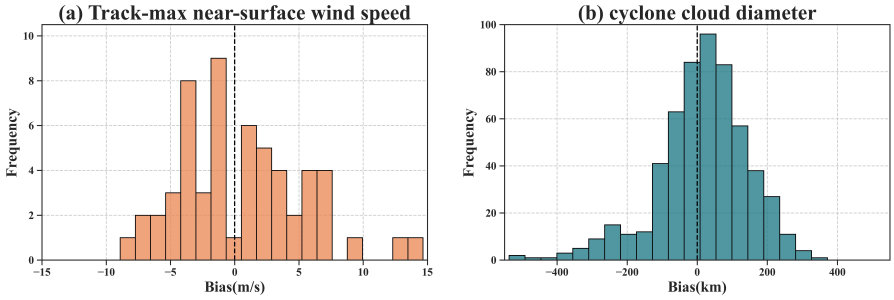


Figure 14: Frequency distribution of bias in (a) Track-max near-surface wind speed and (b) diameter between matched cyclones in the Rojo and IMPMCT datasets (Rojo minus IMPMCT). The cyclone diameter in IMPMCT is calculated as the average of the width and length of the bounding box enclosing the cyclone.

The comprehensiveness of the dataset is constrained by the cyclone representation capabilities of ERA5 reanalysis and the availability of remote sensing data. Since the number of in-orbit satellites carrying the AVHRR sensor peaked around 2013, the IMPMCT track dataset includes the highest number of tracks during this period. Additionally, due to the use of more lenient identification thresholds, IMPMCT tracks typically include longer life compared to the Stoll dataset. The extended portions of these tracks may include: weak vorticity periods during the early/late stages of cyclone development or the vortices pass over land/sea-ice, or re-development processes of vortices after interacting with blocked extropical cyclones or frontal zones. If users require only the core development phases of tracks, they should select segments based on vortex properties that represent the system's core development. The dataset also includes some tracks with high vorticity at their start/end points, which may arise from splitting/merging events or jumps of the vortex center position during tracking.

The dataset does not explicitly distinguish between PMCs and PLs due to the time-sparse wind speed data, particularly when the cyclone's wind speed at a given time step falls below the 15 m s^{-1} threshold. In such cases, it is impossible to determine whether the cyclone is a PMC or merely in a weaker phase of a PL. Further expert analysis using additional criteria is required to differentiate these categories. The VCI images are derived from geographical grids calculated using the Haversine formula, so the representativeness of grid points for actual distances decreases with increasing latitude and distance from the origin. However, given that most PMCs are located below 80° N latitude and have scales smaller than 500 km and distance from grid origin less than 150km, cyclone features extracted from VCI images remain sufficiently accurate. Due to the low resolution of AVHRR infrared images at scan edges, a significant portion of VCI images appear blurred. However, these images are retained as long as cyclonic features remain recognizable, prioritizing the preservation of high temporal resolution for cyclone track records. Additionally, while the YOLOv8-obb-pose model facilitates detection and feature extraction of cyclonic cloud characteristics in VCI images, the process still involves subjective steps to ensure continuity in cyclone features (e.g., size, type, and position). This implies that objective methods for constructing multi-parameter PMC track datasets remain underdeveloped. Consequently, cyclone-evolution-aware deep-learning tracking algorithms could

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

further enhance the efficiency of track construction.

5 Code and data availability

The IMPMCT dataset described in this paper is freely accessible on Zenodo via the following link: https://doi.org/10.5281/zenodo.15355602 (Fang et al., 2025), accompanied by comprehensive documentation. All code is developed in Python and stored at: https://github.com/thebluewind/IMPMCT.

6 Conclusion

The Integrated Multi-source Polar Mesoscale Cyclone Track (IMPMCT) dataset represents a major advancement in the study of polar mesoscale cyclonic systems. By integrating ERA5 reanalysis, AVHRR infrared imagery, and QuikSCAT/ASCAT wind data, this dataset provides a comprehensive record of 1,184 vortex tracks, 16,630 cyclonic cloud features, and 4,373 wind speed observations across the Nordic Seas (2001-2024). This integrated approach overcomes key limitations of previous single-source datasets by enhancing detection sensitivity for weaker polar mesoscale cyclones (PMCs), capturing complete lifecycle evolution from genesis to dissipation, and providing simultaneous cloud morphology and wind fields observations. Rigorous validation against established datasets (Stoll et al., 2022) confirms IMPMCT's accuracy, demonstrating 90 % spatial consistency with track points alignments within 50 km and minimal parameter discrepancies including a 1.09×10^{-5} s^{-1} mean absolute difference in relative vorticity and 0.43 hPa mean absolute difference in sea-level pressure.

IMPMCT could serves as a critical benchmark for evaluating high-latitude climate model performance, a unique case library for comparative studies of PLs and PMCs regarding their formation mechanisms, intensity thresholds, and sea-ice interaction dynamics, and an essential resource for improving polar maritime hazard forecasting. While representing a major step forward, IMPMCT's coverage remains constrained by satellite data availability, highlighting the need for future expansion to pan-Arctic domains as satellite coverage improves and development of more objective cyclone classification algorithms.

Author contributions

RF conceived the experimental design and authored the manuscript. WG contributed to refining the methodologies. XL and HD conducted the research investigations and managed data collection. ZC and CZ contributed to the interpretation of the results. JD and LL provided critical guidance, reviewed, and revised the initial draft. All the authors contributed to the discussions and paper revision.

Competing interests

The contact author has declared that none of the authors has any competing interests.

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

Acknowledgements

The work has been jointly financially by the project of National Key R&D Program of China (Project 2021YFC2802501) and NSF of China (No. 42476205).

References

Andersson, A., Fennig, K., Klepp, C., Bakan, S., Graßl, H., and Schulz, J.: The hamburg ocean atmosphere parameters and fluxes from satellite data – HOAPS-3, Earth Syst. Sci. Data, 2, 215–234, https://doi.org/10.5194/essd-2-215-2010, 2010.

Businger, S. and Steven Businger, R. J. R.: Cyclogenesis in cold air masses, Wea. Forecasting, 4, 133–156, https://doi.org/10.1175/1520-0434(1989)004<0133:cicam>2.0.co;2, 1989.

Bromwich, D. H.: Mesoscale cyclogenesis over the southwestern ross sea linked to strong katabatic winds, Mon. Wea. Rev., 119, 1736–1753, https://doi.org/10.1175/1520-0493(1991)119<1736:MCOTSR>2.0.CO;2, 1991.

Blechschmidt, A. -M.: A 2-year climatology of polar low events over the Nordic Seas from satellite remote sensing, Geophys. Res. Lett., 35, L09815, https://doi.org/10.1029/2008GL033706, 2008.

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.

Condron, A., Bigg, G. R., and Renfrew, I. A.: Polar Mesoscale Cyclones in the Northeast Atlantic: Comparing Climatologies from ERA-40 and Satellite Imagery, Mon. Weather Rev., 134, 1518–1533, https://doi.org/10.1175/MWR3136.1, 2006.

Claud, C., Duchiron, B., and Terray, P.: Associations between large-scale atmospheric circulation and polar low developments over the north atlantic during winter, J. Geophys. Res., 112, D1201, https://doi.org/10.1029/2006JD008251, 2007.

Condron, A. and Renfrew, I. A.: The impact of polar mesoscale storms on northeast atlantic ocean circulation, Nature Geosci, 6, 34–37, https://doi.org/10.1038/ngeo1661, 2013.

Clancy, R., Bitz, C. M., Blanchard-Wrigglesworth, E., McGraw, M. C., and Cavallo, S. M.: A cyclone-centered perspective on the drivers of asymmetric patterns in the atmosphere and sea ice during Arctic cyclones, J. Clim., 1–47, https://doi.org/10.1175/JCLI-D-21-0093.1, 2022.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. -J., Park, B. -K., Peubey, C., De Rosnay, P., Tavolato, C., Thépaut, J. -N., and Vitart, F.: The ERA-interim reanalysis: Configureuration and performance of the data assimilation system, Quart J Royal Meteoro Soc, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Day, J. J., Holland, M. M., and Hodges, K. I.: Seasonal differences in the response of Arctic cyclones to climate change in CESM1, Clim Dyn, 50, 3885–3903, https://doi.org/10.1007/s00382-017-3767-x, 2018.

Fang, R., Ding, J.: IMPMCT: A Dataset of Integrated Multi-source Polar Meso-Cyclone Tracks, Zenodo[data set], https://doi.org/10.5281/zenodo.15355602, 2025

Golubkin, P., Smirnova, J., and Bobylev, L.: Satellite-Derived Spatio-Temporal Distribution and Parameters of North Atlantic Polar Lows for 2015–2017, ATMOSPHERE-BASEL, 12, 224, https://doi.org/10.3390/atmos12020224, 2021.

Gurvich, I., Pichugin, M., and Baranyuk, A.: Satellite multi-sensor data analysis of unusually strong polar lows over the chukchi and beaufort seas in october 2017, Remote Sensing, 15, 120, https://doi.org/10.3390/rs15010120, 2022.

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

Harrold, T. W. and Browning, K. A.: The polar low as a baroclinic disturbance, Quart J Royal Meteoro Soc, 95, 710–723, https://doi.org/10.1002/qj.49709540605, 1969.

Harold, J. M., Bigg, G. R., and Turner, J.: Mesocyclone activity over the north-east atlantic. Part 1: Vortex distribution and variability, Int. J. Climatol., 19, 1187–1204, https://doi.org/10.1002/(SICI)1097-0088(199909)19:11<1187::AID-JOC419>3.0.CO;2-Q, 1999.

Haakenstad, H., Breivik, Ø., Furevik, B. R., Reistad, M., Bohlinger, P., and Aarnes, O. J.: NORA3: A nonhydrostatic high-resolution hindcast of the north sea, the norwegian sea, and the barents sea, J APPL METEOROL CLIM, 60, 1443–1464, https://doi.org/10.1175/JAMC-D-21-0029.1, 2021.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., De Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Quart J Royal Meteoro Soc, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Jocher, G., Chaurasia, A., Qiu, J.: Ultralytics YOLO (Version 8.0.0), GitHub[code] https://github.com/ultralytics/ultralytics, 2023

Kalluri, S., Cao, C., Heidinger, A., Ignatov, A., Key, J., and Smith, T.: The advanced very high resolution radiometer: Contributing to earth observations for over 40 years, Bulletin of the American Meteorological Society, 102, E351–E366, https://doi.org/10.1175/BAMS-D-20-0088.1, 2021.

Laffineur, T., Claud, C., Chaboureau, J.-P., and Noer, G.: Polar lows over the nordic seas: Improved representation in ERA-interim compared to ERA-40 and the impact on downscaled simulations, Mon. Weather Rev., 142, 2271–2289, https://doi.org/10.1175/MWR-D-13-00171.1, 2014.

Mallet, P., Claud, C., Cassou, C., Noer, G., and Kodera, K.: Polar lows over the nordic and labrador seas: Synoptic circulation patterns and associations with north atlantic-europe wintertime weather regimes, JGR Atmospheres, 118, 2455–2472, https://doi.org/10.1002/jgrd.50246, 2013.

Michel, C., Terpstra, A., and Spengler, T.: Polar Mesoscale Cyclone Climatology for the Nordic Seas Based on ERA-Interim, Journal of Climate, 31, 2511–2532, https://doi.org/10.1175/JCLI-D-16-0890.1, 2018.

Noer, G., Saetra, Ø., Lien, T., and Gusdal, Y.: A climatological study of polar lows in the Nordic Seas, Quart J Royal Meteoro Soc, 137, 1762–1772, https://doi.org/10.1002/qj.846, 2011.

Orimolade, A. P., Furevik, B. R., Noer, G., Gudmestad, O. T., and Samelson, R. M.: Waves in polar lows, JGR Oceans, 121, 6470–6481, https://doi.org/10.1002/2016JC012086, 2016.

Parkinson, C. L. and Comiso, J. C.: On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm, Geophys. Res. Lett., 40, 1356–1361, https://doi.org/10.1002/grl.50349, 2013.

Pezza, A., Sadler, K., Uotila, P., Vihma, T., Mesquita, M. D. S., and Reid, P.: Southern hemisphere strong polar mesoscale cyclones in high-resolution datasets, Clim Dyn, 47, 1647–1660, https://doi.org/10.1007/s00382-015-2925-2, 2016.

Rasmussen, E.: An investigation of a polar low with a spiral cloud structure, J. Atmos. Sci., 38, 1785–1792, https://doi.org/10.1175/1520-0469(1981)038<1785:aioapl>2.0.co;2, 1981.

Rasmussen, E. A. and Turner, J. (Eds.): Polar Lows: Mesoscale Weather Systems in the Polar Regions, Cambridge University Press, Cambridge, UK; New York, 612 pp., 978-0-521-62430-5, 2003.

Rojo, M., Claud, C., Mallet, P.-E., Noer, G., Carleton, A. M., and Vicomte, M.: Polar low tracks over the Nordic Seas: a 14-winter climatic analysis, TELLUS A, 67, 24660, https://doi.org/10.3402/tellusa.v67.24660, 2015.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A.: You only look once: Unified, real-time object detection, arXiv [preprint], http://arxiv.org/abs/1506.02640, 9 May 2016.

https://doi.org/10.5194/essd-2025-186
Preprint. Discussion started: 12 May 2025

Rojo, M., Claud, C., Noer, G., and Carleton, A. M.: In Situ Measurements of Surface Winds, Waves, and Sea State in Polar Lows Over the North Atlantic, JGR Atmospheres, 124, 700–718, https://doi.org/10.1029/2017JD028079, 2019.

Shimizu, S. and Uyeda, H.: Algorithm for the identification and tracking of convective cells based on constant and adaptive threshold methods using a new cell-merging and -splitting scheme, J METEOROL SOC JPN, 90, 869–889, https://doi.org/10.2151/jmsj.2012-602, 2012.

Smirnova, J. E., Golubkin, P. A., Bobylev, L. P., Zabolotskikh, E. V., and Chapron, B.: Polar low climatology over the Nordic and Barents seas based on satellite passive microwave data, Geophysical Research Letters, 42, 5603–5609, https://doi.org/10.1002/2015GL063865, 2015.

Smirnova, J. and Golubkin, P.: Comparing polar lows in atmospheric reanalyses: Arctic system reanalysis versus ERA-interim, Mon. Weather Rev., 145, 2375–2383, https://doi.org/10.1175/MWR-D-16-0333.1, 2017.

Stoll, P. J., Graversen, R. G., Noer, G., and Hodges, K.: An objective global climatology of polar lows based on reanalysis data, Quart J Royal Meteoro Soc, 144, 2099–2117, https://doi.org/10.1002/qj.3309, 2018.

Stoll, P. J., Valkonen, T. M., Graversen, R. G., and Noer, G.: A well-observed polar low analysed with a regional and a global weather-prediction model, Quart J Royal Meteoro Soc, 146, 1740–1767, https://doi.org/10.1002/qj.3764, 2020.

Stoll, P. J., Spengler, T., Terpstra, A., and Graversen, R. G.: Polar lows – moist-baroclinic cyclones developing in four different vertical wind shear environments, Polar lows, 2021.

Smedsrud, L. H., Muilwijk, M., Brakstad, A., Madonna, E., Lauvset, S. K., Spensberger, C., Born, A., Eldevik, T., Drange, H., Jeansson, E., Li, C., Olsen, A., Skagseth, Ø., Slater, D. A., Straneo, F., Våge, K., and Årthun, M.: Nordic seas heat loss, atlantic inflow, and arctic sea ice cover over the last century, Reviews of Geophysics, 60, e2020RG000725, https://doi.org/10.1029/2020RG000725, 2022.

Stoll, P. J.: A global climatology of polar lows investigated for local differences and wind-shear environments, Weather Clim. Dynam., 3, 483–504, https://doi.org/10.5194/wcd-3-483-2022, 2022.

Terpstra, A., Michel, C., and Spengler, T.: Forward and Reverse Shear Environments during Polar Low Genesis over the Northeast Atlantic, Monthly Weather Review, 144, 1341–1354, https://doi.org/10.1175/MWR-D-15-0314.1, 2016.

Verezemskaya, P., Tilinina, N., Gulev, S., Renfrew, I. A., and Lazzara, M.: Southern ocean mesocyclones and polar lows from manually tracked satellite mosaics, GEOPHYS RES LETT, 44, 7985–7993, https://doi.org/10.1002/2017GL074053, 2017.

Wilhelmsen, K.: Climatological study of gale-producing polar lows near norway, Tellus A, 37A, 451–459, https://doi.org/10.1111/j.1600-0870.1985.tb00443.x, 1985.

Watanabe, S. I., Niino, H., and Yanase, W.: Climatology of Polar Mesocyclones over the Sea of Japan Using a New Objective Tracking Method, Mon. Weather Rev., 144, 2503–2515, https://doi.org/10.1175/MWR-D-15-0349.1, 2016.

Yanase, W., Niino, H., Watanabe, S. I., Hodges, K., Zahn, M., Spengler, T., and Gurvich, I. A.: Climatology of polar lows over the sea of japan using the JRA-55 reanalysis, Journal of Climate, 29, 419–437, https://doi.org/10.1175/JCLI-D-15-0291.1, 2016.

Yan, Z., Wang, Z., Peng, M., and Ge, X.: Polar Low Motion and Track Characteristics over the North Atlantic, Journal of Climate, 36, 4559–4569, https://doi.org/10.1175/JCLI-D-22-0547.1, 2023.

Zappa, G., Shaffrey, L., and Hodges, K.: Can Polar Lows be Objectively Identified and Tracked in the ECMWF Operational Analysis and the ERA-Interim Reanalysis?, Mon. Weather Rev., 142, 2596–2608, https://doi.org/10.1175/MWR-D-14-00064.1, 2014.