An evaluation of IMERG and ERA5 quantitative precipitation estimates over the Southern Ocean using shipborne observations

E. Montoya Duque[a,b] , Y. Huang[a,b] , P.T. May[c] , S.T. Siems[b,c,d]

[a] *The University of Melbourne, Melbourne, Australia*

[b] *Australian Research Council Centre of Excellence for Climate Extremes, Melbourne, Australia*

[c] *Monash University, Melbourne, Australia*

[d] *Securing Antarctica's Environmental Future (SAEF)*

Corresponding author: E. Montoya Duque,emontoyaduqu@student.unimelb.edu.au

Early Online Release: This preliminary version has been accepted for publication in *Journal of Applied Meteorology and Climatology*, may be fully cited, and has been assigned DOI 10.1175/JAMC-D-23-0039.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

ABSTRACT: Recent voyages of the Australian RV Investigator across the remote Southern Ocean have provided unprecedented observations of precipitation made with both an OceanRAIN maritime disdrometer and a dual-polarization C-band weather radar (OceanPOL). This present study employs these observations to evaluate the Global Precipitation Mission (GPM) Integrated Multi-satellitE Retrievals (IMERG) and the ECMWF reanalysis (ERA5) precipitation products. Working at a resolution of 60 minutes and 0.25° (~25 km), light rain and drizzle are most frequently observed across the region. The IMERG product overestimated precipitation intensity when evaluated against the OceanRAIN but captured the frequency of occurrence well. Looking at the synoptic/process scale, IMERG was found to be the least accurate (overestimated intensity) under warm frontal and high latitude cyclone conditions, where multi-layer clouds were commonly present. Under post-frontal conditions, IMERG underestimates the precipitation frequency. In comparison, ERA5's skill was more consistent across various synoptic conditions, except for high-pressure conditions where the precipitation frequency (intensity) was highly overestimated (underestimated). Using the OceanPOL radar, an area-to-area analysis (fractional skill score) finds that ERA5 has greater skill than the IMERG. There is little agreement in the phase classification between the OceanRAIN disdrometer, IMERG, and ERA5. The comparisons are complicated by the various assumptions for phase classification in the different datasets.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

SIGNIFICANCE STATEMENT: Our best quantitative estimates of precipitation over the remote, pristine Southern Ocean (SO) continue to suffer from a high degree of uncertainty, with large differences present among satellite-based and reanalysis products. New instrumentation on the RV Investigator, specifically a dual-polarization C-band radar (OceanPOL) and a maritime disdrometer (OceanRain), provide unprecedented high-quality observations of precipitation across the SO that will aid into improving precipitation estimates in this region. We use these observations to evaluate the IMERG and ERA5 precipitation products. We find that, in general, IMERG overestimates precipitation intensity, but captures the frequency of occurrence well. In comparison, ERA5 was found to overestimate the frequency of precipitation. Using the OceanPOL radar, an area-to-area analysis finds that ERA5 has greater skill than the IMERG.

1. Introduction

An accurate representation of precipitation is necessary for an integrated understanding of the Earth's climate system. Yet large uncertainties in precipitation estimates exist over the Southern Ocean (SO) (Boisvert et al. 2020; Manton et al. 2020), which is a challenging region because of sparse surface sites, large sea ice masses, and strong winds and waves (Burdanowitz et al. 2019; Boisvert et al. 2020; Siems et al. 2022). This has hindered our ability to understand a wide range of climate and atmospheric processes in this region, as well as their far-reaching impacts through teleconnections (Bodas-Salcedo et al. 2014; Ceppi et al. 2016; Gettelman et al. 2019).

Given the scarcity of in-situ data, satellites have been the foremost tools for detecting and estimating precipitation over remote regions. In fact, a global estimation of precipitation is only possible via satellites, due to their ability to provide near-global and near-continuous coverage (Bumke et al. 2016). However, satellite-based products are still only indirect measures of precipitation using passive and/or active remote sensing instruments (Huffman et al. 2019a). The Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (IMERG) product is currently one of the most advanced satellite precipitation products available, calibrating passive retrievals from across the GPM constellation against active Ku-band radar retrievals from the GPM Core Observatory satellite (Skofronick-Jackson et al. 2017). Products such as these have predominately been developed and calibrated using observational data over the Northern Hemisphere (NH), making them

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

prone to large uncertainties over the pristine SO (Behrangi and Song 2020), where the dynamical and microphysical features of the precipitation systems differ (Huang et al. 2015).

Reanalysis products, such as ERA5 (Hersbach et al. 2020), are an attractive alternative source of precipitation estimates. Unlike satellite-based products, reanalysis precipitation is derived from model simulations, whose accuracy is constrained by model dynamics, model physics, and assimilated information (Hersbach et al. 2020). Small-scale and local processes, such as convection, are often missed or poorly represented due to the limited model resolutions and various assumptions used in the parametrization schemes (Behrangi et al. 2016).

The large discrepancies between the reanalysis and satellite-based precipitation estimates over the SO have translated into inconsistencies in precipitation trends and thermodynamic phase classifications (Behrangi et al. 2014, 2016; Boisvert et al. 2020; Manton et al. 2020). Manton et al. (2020) showed that trends of precipitation monthly anomalies derived from several widely used products vary, even with opposite signs over different sectors of the SO. Snowfall precipitation intensity estimates between reanalysis and satellite products also diverge, especially at mid and high latitudes of the SO (Boisvert et al. 2020). These challenges necessitate a comprehensive assessment of the reanalysis and satellite products across the range of SO precipitation mechanisms (Khan and Maggioni 2019).

Precipitation evaluation studies using surface measurements over the SO are often limited to a few island sites, with the records collected from Macquarie Island (MAC) being of exceptional value given both the quality and extent, dating back to 1948. An evaluation of ERA-Interim product using MAC rain gauge data showed that ERA-Interim underestimated the annual precipitation from MAC by about 6.8% (Wang et al. 2015). It was also found that ERA-Interim over-estimated the precipitation in the mid-latitude cyclones and fronts and underrepresented the precipitation away from these systems, errors that were largely fixed by ERA5 (Lang et al. 2018; Wang et al. 2015; Boisvert et al. 2020). More recently, Tansey et al. (2022) used a blended set of in-situ and remote precipitation estimates to evaluate the CloudSat precipitation products, finding that in the MAC area CloudSat estimates tend to underestimate light liquid precipitation frequency and intensity, and mixed-phase precipitation frequency. Despite these efforts, a more systematic assessment of the precipitation products, under the varying meteorology across the broader extent of the SO, is largely absent.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

While ship-based data from Research Vessels (RVs) and ocean buoys represent another useful set of measurements, the rain gauges of standard design being conventionally employed can often be biased by sea spray, intense winds, and spatial and temporal coverage (Klepp et al. 2010, 2020). The Ocean Rainfall And Ice-phase precipitation measurement Network (OceanRAIN) disdrometer, which was specifically designed to reduce uncertainties arising from rough oceanic and atmospheric conditions, offers much greater confidence in ship-based observations (Klepp et al. 2010). Initiated in 2009, the OceanRAIN project is arguably the only systematic long-term shipboard precipitation data collection effort to date (Klepp et al. 2020; Klepp 2020). The observations have been used to evaluate a variety of satellite and reanalysis surface precipitation products, with a heavy focus on the NH oceans and bulk precipitation statistics (Bumke et al. 2012, 2016; Burdanowitz et al. 2019). Since the first deployment over the SO in 2016, OceanRAIN has provided the scientific community with unique and high-quality precipitation information in this region. Using approximately 790 hours of OceanRAIN disdrometer observations, Montoya Duque et al. (2022) found that precipitation over the SO is dominated by very light and light precipitation (below 1 mm hr^{-1}) and that the high latitude region commonly has mixed-phase and snow precipitation.

Another significant advance in recent years has been the availability of a dual-polarization C-band precipitation radar (OceanPOL) over the SO. Onboard the RV Investigator (operated by the Australian Marine National Facility), the OceanPOL commenced its mission in 2018, making extensive and near-continuous measurements across the Australian sector of the SO through multiple cruises. Since its deployment, the OceanPOL has produced a wealth of fresh precipitation details with unprecedented spatio-temporal coverage, making the dataset useful for evaluation purposes.

Motivated by these emerging opportunities, this research aims to perform a systematic evaluation of two widely used precipitation products over the SO, IMERG and ERA5, using the OceanRAIN disdrometer and OceanPOL radar observations collected from multiple recent field campaigns. We also extend the analysis to examine precipitation estimates under various synoptic conditions, using the recently established clustering synoptic classification by Truong et al. (2020) and Montoya Duque et al. (2022). The research is guided by two questions: How accurate are precipitation estimates over the Australian sector of the Southern Ocean compared to ship-based measurements? How does the accuracy vary under the various synoptic conditions? The data and methods are

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

described in section 2, followed by the results in sections 3 to 6, discussion, and concluding remarks in section 7.

2. Data and methods

a. Data

1) Surface based and remote-sensing measurements

The Research Vessel (RV) Investigator (CSIRO 2022) has deployed an OceanRAIN disdrometer (herein ODM470) (Klepp et al. 2018; Protat and CSIRO 2020) since 2016 and a C-band dual-polarization weather radar (OceanPOL) since 2018 (Louf et al. 2019) over the SO and Australian surrounding (Table 1). A total of six research voyages (181 days) have been undertaken over the Australian sector of the SO and the Antarctic coast between the Austral summer season or early fall 2016 and 2018 (Figure 1). Two of the voyages were part of the CAPRICORN (Clouds, Aerosols, Precipitation, Radiation, and atmospherIc Composition Over the SoutheRn OceaN) field campaigns (McFarquhar et al. 2021), where the observations were enhanced by a suite of cloud-aerosol-radiation measurements such as a W-band cloud radar (Table 1), providing further insights in the case analysis presented in section 3.

Table 1. Remote-sensing and in-situ instruments deployed on the RV Investigator that are employed in this study

Instrument	Resolution/coverage	Variables	Description	References
BASTA cloud radar	1 min. Coverage: two case studies (around three days each, Section 3)	Reflectivity	The W-band cloud radar (94.95 GHz) was mounted on a stabilized platform with four vertical resolutions (12.5, 25,100, 200 m), a valid signal between 40 m and 12 km, and detectable radar reflectivity –45 dBZ at 1 km	Delanoë et al. (2016)
OceanPOL C-band weather radar	5 minute volume scans. Coverage two voyages: i) 2018-01-11 to 2018-02-22 and ii) 2018-03-03 to 2018-03-2	Reflectivity, Differential reflectivity, precipitation rate estimate, Microphysical classification, cross-correlation	C-band dual-polarization radar (5.6 GHz) with a 1.3 degree beamwidth that uses a pulse rate of 1000Hz with a 1 μs pulse and 125 m range sampling. Eleven tilts were used in the scanning, with a resulting resolution of 8 km height, and a 150 km radius	Louf et al. (2019) Louf and Protat (2020)
OceanRAIN disdrometer ODM470	1 minute. Coverage six voyages, see Figure 1	Precipitation rate, phase classification	The ODM470 emits a light from a diode at 880 nm, measuring a volume of 120 mm in length and 22 mm in diameter. The particles are sorted in 128 bins with logarithmic increase from 0.04 to 22 mm in diameter. An anemometer and precipitation detector are also included	Klepp et al. (2010) Klepp et al. (2018)

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

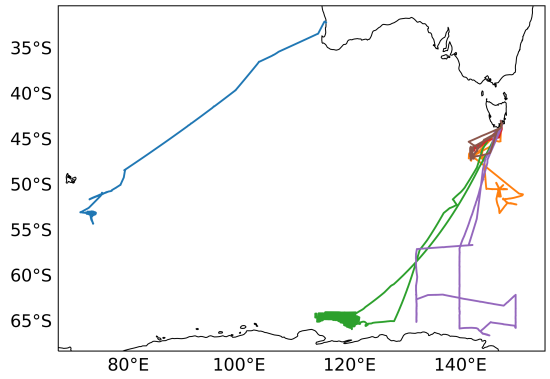


FIG. 1. Trajectories of research voyages undertaken over the Southern Ocean (SO) by the RV Investigator between January 2016 and March 2018. Blue 2016-01-07 to 2016-02-07. Orange 2016-03-14 to 2016-04-16 (CAPRICORN I). Green 2017-01-14 to 2017-03-05. Red 2017-03-17 to 2017-03-27. Purple 2018-01-11 to 2018-02-22 (CAPRICORN II). Brown 2018-03-03 to 2018-03-21.

Compared to other existing disdrometers that are not designed for shipboard operation, the ODM470 was specially developed to meet all-weather shipboard requirements and has been shown to have improved performance under strong wind conditions (Klepp et al. 2018). In this study, we use the OceanRAIN v2.0, which provides additional ship records and radar-related variables that were not available in previous versions (Klepp 2020). However, quantitative measurements of snow and mixed-phase precipitation are still challenging with the ODM470 and involve large uncertainty, especially during strong precipitation events. The liquid-to-solid ratio from mixed-phase events remains poorly quantified, and snow shape and density assumptions do not necessarily represent the snow variability (Klepp et al. 2010, 2018). After removing obvious outliers from the ODM470 precipitation rates, there remains a small number of samples with suspiciously large intensities (above 50 mm hr^{-1}), which are related to one precipitation event that occurred at high latitudes during mixed-phase periods. These data should be treated with greater caution. The hourly thermodynamic phase is estimated based on the hourly frequency of each phase. The liquid or snow phase is defined when the respective frequency of occurrence is at least 90%; mixed phase is assigned when liquid or snow is between 10 and 90%.

OceanPOL weather radar product v2020, provided by the Australian Bureau of Meteorology (BoM), include estimates of precipitation rate produced from the dual-polarization rainfall algorithm (see Thompson et al. (2018)). The algorithm employs information from the horizontal radar

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

reflectivity and differential reflectivity. The cross-correlation coefficient, ρ(HV(0)), is used for quality control since it indicates the presence of non-Rayleigh scatters, the similarity between scatters' phase, and orientation. Compared to single-polarization radars, dual-polarization information allows for better detection of the melting layer, quantitative rainfall estimates, and precipitation type classification (Thompson et al. 2015). The Thompson et al. (2018) algorithm was calibrated for liquid precipitation using data from two tropical islands, which may not be optimal for the SO environment. Additionally, a sensitivity analysis of the OceanPOL radar histograms suggests that at a long range (greater than about 100 km), reflectivities below 10 dBZ are not measured, thus limiting the skill of rainfall detection and estimates for drizzle/light precipitation events. To avoid sea-clutter contamination, we use the gridded precipitation rate estimates produced at an altitude of 500 m, the lowest altitude representing near-surface precipitation. Further, we exclude all values within a 5 km radius from the radar location before interpolating the data onto a 0.25° x 0.25° grid using the nearest grid-point method. In the operational dataset, rainfall (i.e. liquid precipitation) estimates were produced for all valid radar grid points regardless of the precipitation phases. Thus, we identify rain grids based on the microphysical classifier (Thompson et al. 2014), a ρ(HV(0)) value above 0.9, and when the corresponding hourly ODM470 phase classification indicates either liquid or no precipitation at the ship location.

2) Satellite data

The satellite-based estimates of precipitation from the Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (IMERG) product are examined in this study. The IMERG dataset is currently the most advanced remote-sensing precipitation product with a global coverage (Huffman et al. 2019a). The IMERG algorithm intercalibrates, merges, and interpolates microwave, microwave-calibrated infrared, precipitation gauge data, and other data from suitable spaceborne sensors to produce a gridded dataset with a spatio-temporal resolution of 0.1° and 30-min (Huffman et al. 2019a, 2021). In the IMERG product, precipitation estimates can come from Passive MicroWave-only overpasses (PMW), InfraRed-only (IR) overpasses, interpolated values using a morphing-only scheme based on MERRA-2 reanalysis, or an interpolation between PMW-IR or PMW-morphing (herein defined as mixture; see supplementary material) (Li et al. 2022). The IMERG precipitation algorithm is an upgrade from the Tropical Rainfall Measuring

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

Mission (TRMM) product, making IMERG largely tuned for tropical rainfall (Huffman et al. 2019a; Protat et al. 2019b). Unlike TRMM, however, IMERG provides information at higher latitudes (full coverage between $\pm 60^o$ and partial coverage poleward to this latitude), precipitation phase classification, and improved sensitivity to light rain. Field campaigns such as CAPRICORN II had aimed at improving GPM precipitation products, among others. When available, the OceanRAIN product provides freshwater flux data employed by IMERG to improve the ocean precipitation estimates (Klepp et al. 2020; McFarquhar et al. 2021). In this study, we use the rain gauge calibrated precipitation rate estimates (precipCal) and the probability of liquid precipitation (%) for phase classification by using similar thresholds as for the disdrometer data (see Section 2.a.1). The final run of the Level 3 product (Level 3 V06B) includes all satellite overpasses, providing a complete picture of precipitation around the globe (Huffman et al. 2019a). This is especially useful for areas where not a lot of ground-based data is available, such as the SO.

The Himawari-8 satellites, operated by the Japan Meteorological Agency (JMA), carries the Advanced Himawari Imager (AHI) instrument which is a 16-channel multispectral imager capturing visible and infrared images. In section 3 we supplement the precipitation analysis with the Brightness Temperature (BT) as seen from the Himawari-8 AHI channel 14 (11.2 μm, infrared), which has an original spatial resolution of 2 km at nadir (Bessho et al. 2016).

3) Reanalysis data

ERA5 is the fifth generation European Centre for Medium-Range Weather Forecast (ECMWF) atmospheric reanalysis of the global weather, providing estimates of a wide range of atmospheric, land, and oceanic climate variables. The precipitation product includes estimates of precipitation intensity and precipitation classification, generated at 60 min temporal and 0.25° horizontal resolutions (Hersbach et al. 2018, 2020). In this study we use the mean total precipitation rate (mtpr) and the precipitation classification variables presented in Code Table 4.201 from ECMWF. For comparison purposes, we have re-grouped the precipitation classes as follows: liquid (rain, thunderstorm, freezing rain, drizzle, and freezing drizzle), mixed (mixed/ice, wet snow, mixture of rain and snow), and snow (snow, ice pellets, graupel, and hail).

The reanalysis dataset only directly assimilate radar/rain gauge data over the eastern United States of America (US) since 2009, leaving the oceanic and most of land precipitation products as model

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

outputs (Hersbach et al. 2020; Lavers et al. 2022). Nevertheless, previous studies have reported that the hydrological cycle is better represented in ERA5 than in ERA-Interim due to improved data assimilation of indirect measurements such as those from microwave and infrared satellite sensors (e.g., GOES, AVHRR, and Himarawri-8) and improvements in the model's microphysical parameterizations (Hersbach et al. 2020). ERA5 product has been found to produce the most accurate estimates of the mean and seasonal cycle of precipitation over the SO (Boisvert et al. 2020; Manton et al. 2020), but large uncertainties remain in the phase classification and snowfall estimates (Boisvert et al. 2020).

b. Methods

1) Data collocation and averaging

The spatio-temporal geometry and resolution of the OceanRAIN disdrometer, OceanPOL radar, ERA5, and IMERG products are different, confounding any direct comparison, especially when comparing point measurements with area averages (the so-called 'point-to-area' problem) (Loew et al. 2017). To minimize potential errors associated with these issues, we performed a space–time averaging of the disdrometer measurements assuming that the space–time conversion is sufficient to represent the mean precipitation features within the corresponding satellite/reanalysis grid. The hourly mean wind speed measured at the ship is approximately 9.7 m s^{-1}, which corresponds to a distance of about 34 km in an hour. Thus the 0.25° gridded data is comparable to the hourly averaged shipborne measurements on the horizontal scale. Gridded data (ERA5 and IMERG) are collocated at the ship location using the nearest-neighbor(s) method. If the ship trajectory covered more than one grid point (or parts of them) within an hour, the ERA5 and IMERG data are then averaged over those grid points. To test the sensitivity of the results to the space–time window size, time averages over 30 min and 180 min were examined at 0.1° and 0.75° grid resolution. We also tested the collocation method by selecting only the nearest neighbor, as well as a 3 x 3 grid averaging around the mean position of the ship. The results are all found to be largely insensitive to the different spatio-temporal collocation methods used. It should be noted that about 18% (6%) of the ERA5 (IMERG) collocated precipitation rates are below 0.1 mm day^{-1} (i.e., trace precipitation threshold as defined by the World Meteorological Organization WMO, Boisvert et al. (2020)). In practice, we are unable to determine the presence and nature of these very light precipitation events

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

suggested by the gridded products, due to the limited measurement sensitivity of the shipborne instruments. Thus, all our analyses are performed for precipitation intensities above the trace precipitation threshold.

2) Spatial analysis

A common shortcoming associated with point-to-area comparisons is the so-called 'double penalty' problem, i.e., a mismatch with truth gives both a miss and a false alarm (Keil and Craig 2009). To address this problem, spatial verification metrics are often used to compare precipitation products of different resolutions against a common spatial truth. The fraction skill score (FSS), defined in equation 1 (Roberts and Lean 2008), is one of these metrics that directly compare the fractional coverage of grid-mean precipitation (that exceeds a threshold) from two datasets in increasing spatial windows. The FSS returns a value in the range of 0 (no skill) and 1 (perfect match). The useful threshold will determine the minimum useful scale, which is determined by the fractional rainfall over the domain (f, wet-area ratio) of the OceanPOL, if f is small enough, the useful threshold is approximately 0.5 (Mittermaier and Roberts 2010).

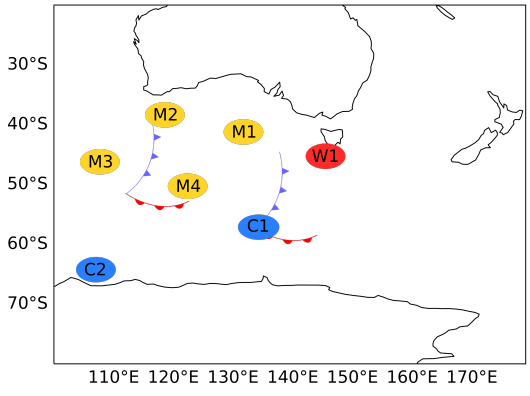
$$
\begin{aligned}
FSS &= 1 - \frac{\frac{1}{N}\sum_N (P_m - P_o)^2}{\frac{1}{N}\left[\sum_N P_m^2 + \sum_N P_o^2\right]}, \\
P_o, P_m &= \text{precipitation fraction, o for OceanPOL, and m for ERA5 or IMERG,} \\
N &= \text{Number of windows,} \\
FSS_{usefull} &= 0.5 + f/2, \text{ where f is the wet-area ratio}
\end{aligned}
\tag{1}
$$

This study evaluated the FSS within a 7 x 7 grid area for the ERA5 and IMERG datasets. This is the maximum number of grid points that fit within the 150 km range of the OceanPOL observations. Additionally, the ERA5, IMERG, and OceanPOL data are being compared at 60 min, 0.25° x 0.25° resolution for the estimated "liquid" precipitation events, where the OceanPOL rainfall retrievals are deemed most reliable (see Section 2.a.1).

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

3) Synoptic classification

A synoptic identification is performed using the k-means clustering results produced by Truong et al. (2020) with 2,186 soundings collected from four different field campaigns (Figure 2). The classification identified seven unique clusters across the SO using 15 thermodynamic variables: temperature, relative humidity, zonal wind, and meridional wind at 925, 850, and 700 hPa levels, respectively, and pressure, air temperature, and relative humidity at the surface. Since the sounding data were only available from two of the six voyages included in our study (i.e., CAPRICORN I and II), we have opted to use the 'pseudo' soundings from the hourly collocated ERA5 data to classify each hour into one of the seven clusters. Each pseudo sounding was assigned to a cluster with the minimum Euclidean distance to one of the seven k-means centroids. For CAPRICORN I and II, our analysis shows that the classification using the ERA5 soundings agrees with the physical sounding classification 89% of the time (not shown). Truong et al. (2022) compare the 'pseudo' and physical soundings and found that ERA5 has an appreciable level of skill in reproducing the clusters over the SO.



Cluster	Characteristics
W1	Characterised by warm air advection from the Australian continent in pre-frontal regions north of ocean polar front
M1	Typically found in high-pressure systems (and ridging) in large scale subsidence
M2	Located at the north-northeast of the cyclone centre close to cold fronts but removed from the cyclone centre
M3	Represents the post-frontal area of a cold front, and northwest of the cyclone centre
M4	A warm front cluster
C1	Represents the conditions poleward of the ocean polar front near cyclones
C2	Indicates the conditions in the coastal Antarctica area, it can be located southwest of the cyclone centre

Fig. 2. A conceptual illustration of the seven thermodynamic clusters derived in Truong et al. (2020). The red circle indicates the warm cluster (W1), the yellow circle the medium clusters (M1–M4), and the dark blue circle the cold clusters (C1–C2). The red (blue) curve with half-circles (triangles) indicates the warm (cold) front. On the right, the main cluster characteristics are also presented. Adapted from Montoya Duque et al. (2022)

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

3. A comparison between precipitation products: example cases

Two examples of precipitation events encountered during the voyages are presented to illustrate some of the common precipitation characteristics and potential issues for the comparison.

a. A Mid-latitude case

In this case, we examined an event when the RV Investigator was located near 50°S and 140°E between 12 UTC on January 16 and 00 UTC on January 20, 2018. High-level clouds and a low-pressure system were located south of the RV Investigator ahead of the warm sector (Figure 3a-c). The warm sector passage was followed by a cold frontal passage 24 hours later (January 19, 2018); no cyclone was present within a distance of 5^o from the vessel (Figure 3c).

The disdrometer indicated two main precipitation periods at around 12 UTC on January 18 and 12 UTC on January 19, 2018. These events coincide with a local minima in the Mean Sea Level Pressure (MSLP), where the pressure drop is higher during the warm sector than during the last cold frontal passage. After the warm sector passage, the ship crosses a cold front and post-frontal conditions, recording no precipitation at the surface, before finally reaching a second (weaker) cold front with precipitation (Figure 4). Surface air temperature increased during the warm sector passage, followed by a steep drop in temperature and relative humidity. The BASTA cloud radar detected upper-level clouds prior to the warm sector passage and some light precipitating clouds during the warm air advection in pre-frontal areas, W1. Low-level precipitating clouds were present during the two cold frontal passages, and no high-level clouds were visible (Figure 4e). These patterns are consistent with features shown in the Himawari-8 brightness temperature (BT) and OceanPOL radar (Figures 3a-c, j-l). The OceanPOL accumulated precipitation is similar to the ODM470 record, while the IMERG collocated precipitation intensities at the ship location are consistently higher than the disdrometer and weather radar estimates (Figure 4a,b). ERA5 overestimates the accumulated precipitation by about 5 mm (Figure 4b), although the discrepancies are less pronounced than with IMERG, overestimating the accumulated precipitation by 15 mm and the precipitation intensity by about 1.5 mm hr^{-1} during the passage of the second cold front (Figure 4a,b).

Looking at the spatial structures, ERA5 produced a rain band north of the ship during the pre-frontal conditions, around 2018-01-18 01 UTC, which is spatially consistent with the precipitation

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

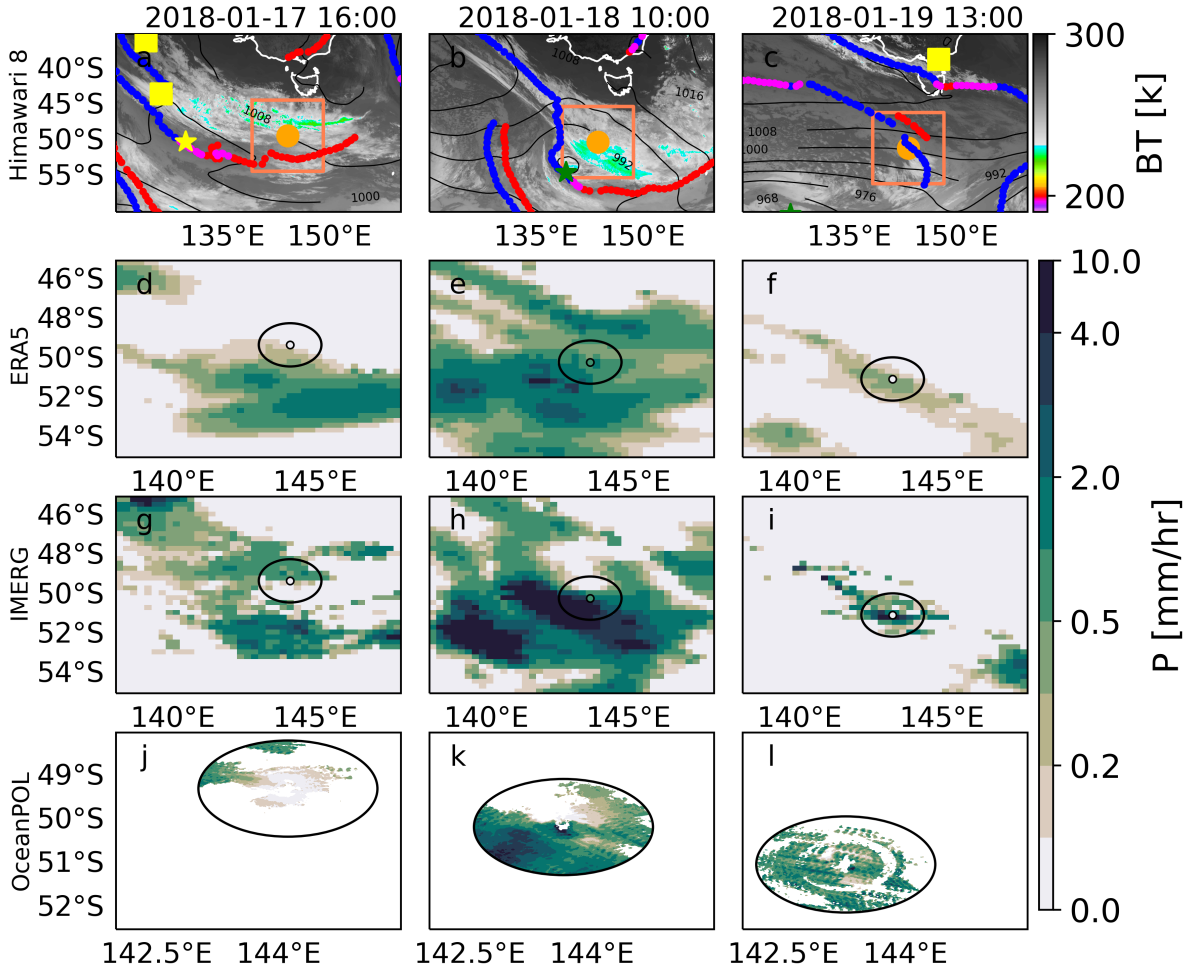


FIG. 3. From left to right, warm air advection in pre-frontal areas at 2018-01-17 16 UTC, warm sector at 2018-01-18 10 UTC, and cold front at 2018-01-19 13 UTC. From top to bottom, the first row shows Himawari-8 Brightness Temperature BT [K] as colored areas, and contours show the Mean Sea Level Pressure (MSLP) from ERA5. The orange dot is the RV Investigator's location framed in the second and third-row visual area. Blue (red/pink) dots are cold (warm/stationary) fronts detected with the Berry et al. (2011) scheme. Squares (stars) represent the open (close) cyclones colored in green (yellow) if they are strong (weak) according to Murray and Simmonds (1991) classification. More details about the front and cyclone detection can also be found in Truong et al. (2020). The second (third) row shows ERA5 (IMERG) hourly precipitation rate at 0.25° spatial resolution, with the OceanPOL coverage area highlighted and the disdrometer precipitation estimate indicated at the center. The last row shows OceanPOL hourly liquid precipitation rate estimate [mm/hr].

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

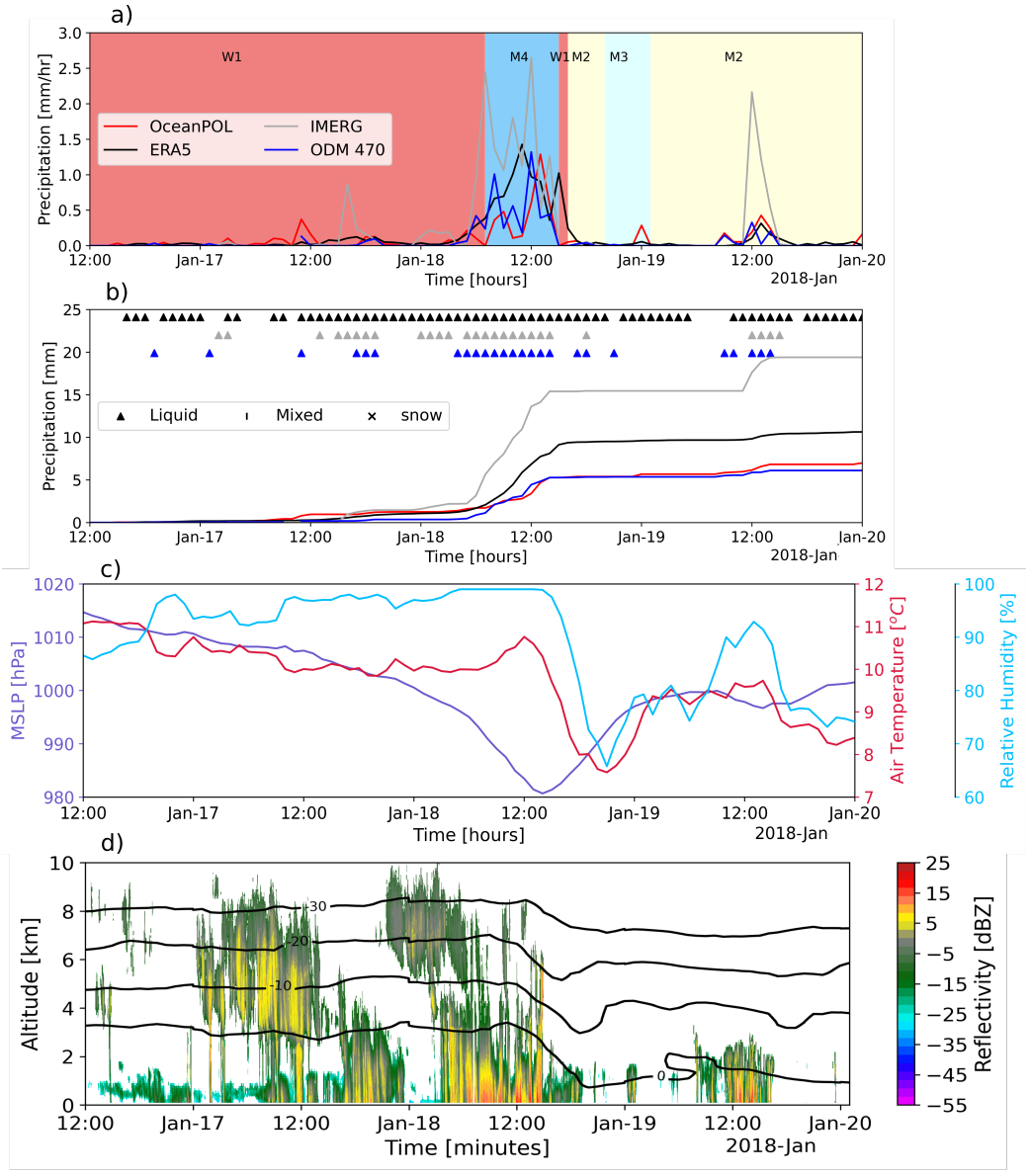


FIG. 4. Time series of a) Precipitation rate [mm/hr] for ERA5 (black), IMERG (gray), ODM470 (blue), and OceanPOL (red); shaded areas indicate the synoptic clusters (Figure 2). b) Accumulated precipitation [mm] for the four datasets (lines) and thermodynamic phase when available (scatters). c) Mean Sea Level Pressure (MSLP, hPa, purple), relative humidity (%, blue), and surface air temperature (°C, red). d) BASTA cloud radar reflectivity (dBZ), ERA5 temperature contours. Between 2018-01-16 12:00 and 2018-01-20 00:00 UTC, the RV Investigator was located near 50°S and 140°E.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

pattern observed by OceanPOL. IMERG showed multiple precipitation bands with the main one located to the south of the ship, which corresponds to the coldest cloud tops depicted in the Himawari-8 BT image (Figure 3). An hour-by-hour analysis (not shown) indicates that IMERG tended to place precipitation in areas where high-level cloud tops were detected with no or light precipitation at the surface. In the absence of cold BT, IMERG captured the spatial pattern of precipitation, as shown by the OceanPOL, reasonably well, but small-scale precipitation features tended to be missed or under-represented. In contrast, ERA5 captured the precipitation patterns reasonably well, although it also tended to miss small-scale precipitation cores or have them with lower intensities. Additional analysis (See the supplementary material for a further explanation) suggests that the main source of IMERG precipitation in this case is associated the (Passive Microwave only) PMW-only (48%) algorithm, especially during the periods when IMERG is seen to overestimate the disdrometer values (Figure 4a). In ERA5, precipitation estimates is primarily the microphysics parametrization (69%).

b. A High-latitude case

In this example, we examined the precipitation structure south of the ocean polar front between 09 UTC on January 23 and 11 UTC on January 27, 2018. The RV Investigator was located near 58°S and 140°E during that time. This period comprises two different high-latitude cyclone passages (C1) near the RV Investigator and a trough passage towards the end of the period. The first cyclone is classified as an 'open/weak' cyclone by the cyclone detection scheme, with low-level clouds observable in the Himawari-8 imagery (Figure 5a-c). The second cyclone is diagnosed as a 'closed/strong' cyclone, with high-altitude clouds and a warm front approaching the vessel from the north. Towards the end (around 06 UTC January 27, 2018), high-level clouds associated with the trough conditions (misclassified as warm front) are present near the vessel, whereas an 'open/weak' cyclone is identified to the east of the ship.

The disdrometer precipitation time series showed three precipitation episodes associated with the two high-latitude cyclones and the trough passage (Figure 6), with recorded precipitation intensities below 0.7 mm hr^{-1} most of the time. The first C1 episode was characterized by low-level clouds below 2 km, while the second C1 episode (around 2018-01-25 12 UTC) had a well-defined cloud band present up to 6 km (Figure 6e). Note that snow and mixed-phase precipitation is likely present

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

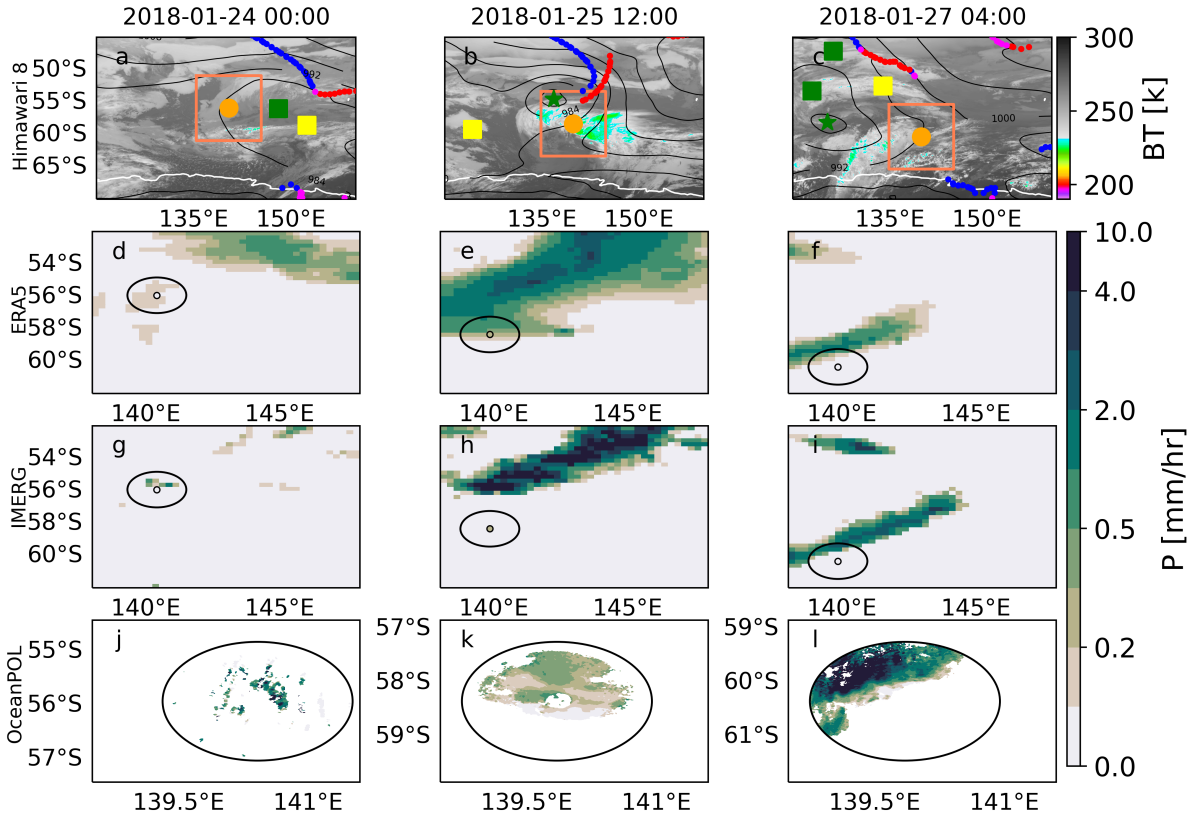


FIG. 5. Same as 3 but for the high-latitude cyclone center example (2018-01-25 13 UTC)

in the vicinity of the ship during the second C1 episode as highlighted in Mace et al. (2022). Clouds up to around 4 km were observed during the third episode under the trough conditions.

The collocated ERA5 precipitation frequently showed higher precipitation intensities (and accumulated values) than the ODM470 records, while IMERG reported lower intensities throughout most of the period, except around 2018-01-27 04 to 06 UTC, where IMERG overestimated the intensity by about 2 mm hr^{-1} compared to ODM470 (Figures 6a-b). For the above-mentioned period, IMERG and OceanPOL reported similar intensities. However, the BASTA cloud radar reflectivities during this period were similar to/smaller than those associated with the second high-latitude cyclone (Figure 6e), suggesting that precipitation characteristics were comparable to that during the second cyclone passage. This is consistent with the similar disdrometer precipitation rate estimates in both periods. There is a poor agreement in the precipitation phase between the different datasets. We note that the ambient surface temperature was below 5°C throughout the

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

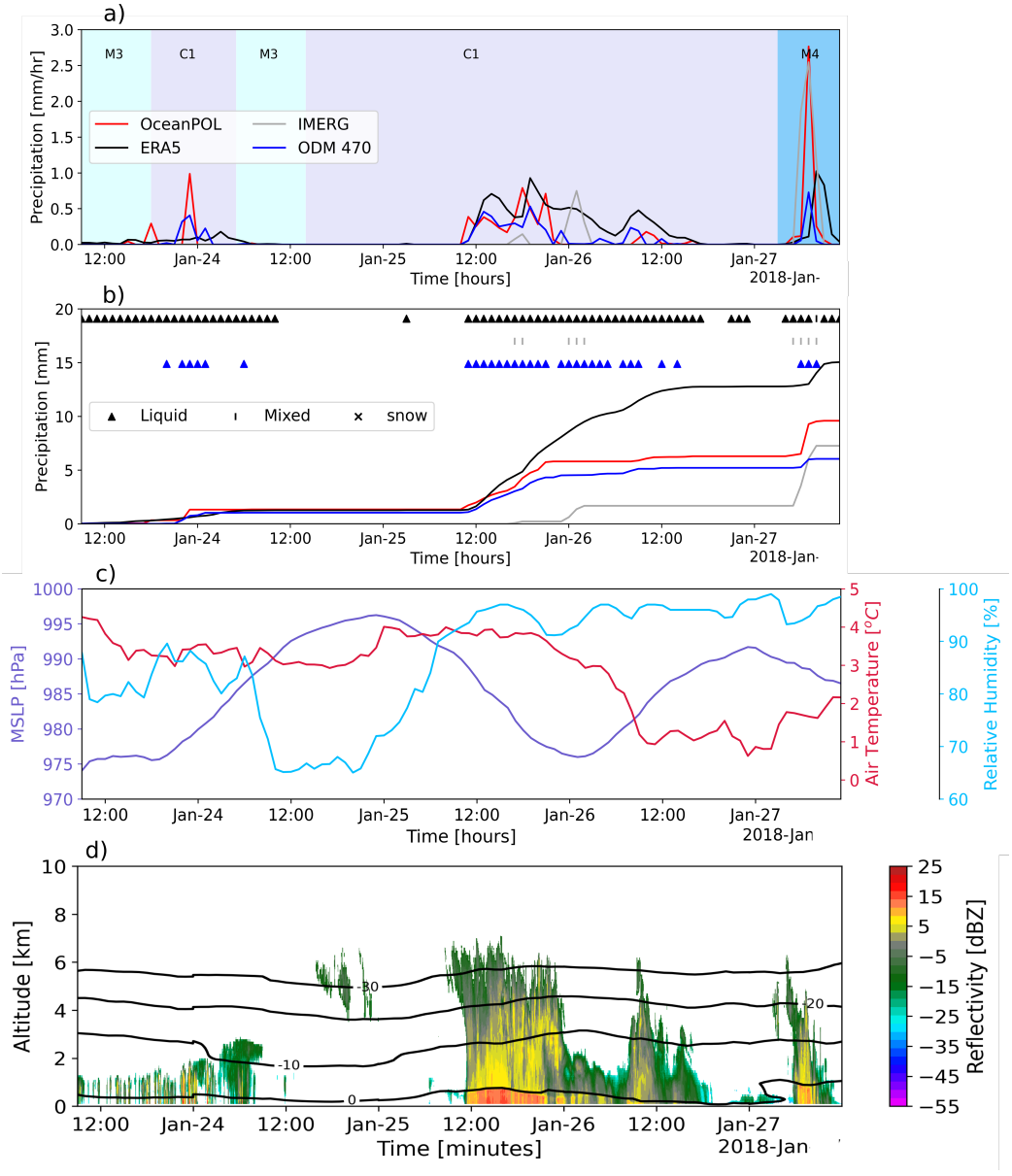


FIG. 6. Same as 4. Between 2018-01-23 09:00 to 2018-01-27 11:00 UTC. The RV Investigator was located near 58°S and 140°E.

entire period, where particles may be partially or wholly melted, resulting in large uncertainty in phase classification (and the resulting precipitation intensity estimates) from all datasets.

Spatially, both ERA5 and IMERG reported a rainband to the north of the vessel under the second high-latitude cyclone center conditions (around 2018-01-25 13 UTC), coincident with high-altitude clouds evident in the cold Himawari-8 BTs (about -53°C). However, the precipitation

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

spatial extent in IMERG seems to be more contracted compared to the ERA5 counterpart, whose spatial precipitation pattern compares better with the OceanPOL radar observations (Figure 5). Similar spatial features can also be found in an hour-by-hour visual analysis of these products (not shown). With the presence of low BT (around -50 °C) from deep clouds at the ship's location (up to 6 km), IMERG tended to produce heavier precipitation as compared to OceanPOL and the disdrometer estimates. When no deep clouds were present, however, IMERG produced reasonable precipitation intensity estimates but a less accurate spatial extension when compared against OceanPOL, especially during the second high-latitude cyclone passage. The morphing-only algorithm heavily influencing the IMERG precipitation estimates (44%) in this case. During the trough conditions, the highest overestimation in precipitation rate is seen and the PMW-only is the primary source of estimate. At 2018-01-25 12 UTC, the precipitation estimates in the rainband are produced by the morphing-only scheme. As in the mid-latitude case, ERA5 precipitation is dominated by the microphysics parametrization 68% of the time (See the supplementary material for a further explanation).

Within these two examples, some discrepancies between the OceanRAIN, OceanPOL, ERA5, and IMERG precipitation estimates are noted, e.g., IMERG precipitation intensity tended to be higher than ODM470 values during periods of low BTs (high cloud tops). ERA5 precipitation patterns compared better with the OceanPOL observations, even for some of the small-scale precipitation events. Indeed, these discrepancies are repeatedly observed in several other cases we have examined (not shown), which necessitates a more comprehensive analysis where ERA5 and IMERG performance can be more systematically evaluated.

4. Statistical evaluation of the precipitation estimates from six voyages

A total of 4346 hours of ODM470 precipitation observations were collected over the six SO voyages conducted between January 2016 and March 2018. Table 2 presents the Frequency of Occurrence (FoO) and accumulated (accum) precipitation for the total precipitation rate and by the thermodynamic phase of collocated values (rows 1-3). The IMERG product records precipitation 22.3% of the time, roughly equal to that of the ODM470 (25.4%), while ERA5, greatly overestimated the FoO at 64.1%. For the accumulated precipitation, however, ERA5 (436 mm, underestimation of 2.3%) is closer to the observed total (448 mm) than IMERG (615 mm,

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

overestimation of 37%). An overestimation by a factor of four is noted for the IMERG mixed-phase precipitation (149 mm to 39 mm), which is the highest among all phases.

TABLE 2. Frequency of occurrence (FoO %) and accumulation (Acumm mm) of total precipitation and by thermodynamic phase for IMERG, ERA5 and the ODM470 (rows 1-3) for the 4346 hours of collocated observations. Concurrent observations of the IMERG and ODM470 (IO) for any precipitation (row 4) and by phase (row 5). Similarly, concurrent observations of ERA5 and ODM470 (EO) (rows 6 and 7) and all three products (IEO) (rows 8 and 9).

	Total Precipitation		Liquid phase		Mixed phase		Snow phase	
	FoO (%)	Accum (mm)	FoO (%)	Accum (mm)	FoO (%)	Accum (mm)	FoO (%)	Accum (mm)
ODM470	25.4	448_O	19.6	367	2.9	39	2.9	42
IMERG	22.3	615_I	11.9	401	7.2	149	3.2	65
ERA5	64.1	436_E	49.1	344	9.0	78	5.9	15
IO_{precip}	12.2	470_I	8.0	371	3.6	80	0.7	18
		326_O	10.1	280	0.7	17	1.3	30
IO_{phase}	9.2	399_I	8.0	371	0.7	10	0.6	17
		202_O		184		16		2
EO_{precip}	24.0	334_E	19.1	260	3.4	67	1.4	6
		441_O	18.6	362	2.7	37	2.7	42
EO_{phase}	20.6	287_E	17.9	253	1.4	28	1.4	6
		329_O		295		26		7
IEO_{precip}	11.8	464_I	7.7	366	3.5	80	0.6	18
		246_E	9.7	201	1.4	44	0.7	2
		323_O	9.8	277	0.7	15	1.3	30
IEO_{phase}	8.6	385_I	7.7	365	0.4	9	0.4	10
		189_E		171		16		1
		197_O		182		13		1

Focusing only on the hours when IMERG and ODM470 observe any precipitation concurrently (IO_{precip}), we again see that IMERG overestimates the accumulated precipitation (470 to 326 mm), most notably for mixed phase (80 to 17 mm). Examining only the hours when the phase of the precipitation agrees for the IMERG and ODM470 products phase (IO_{phase}), we limit the analysis to only 9.2% of the 4346 hours. Perhaps surprisingly, when mixed phase is observed by both datasets, the ODM470 records a greater accumulation of precipitation (16 to 10 mm), although this is for a very small sample size. The same analysis is made for concurrent ERA5 and ODM470 (EO) observations (rows 6,7). When limited to the total precipitation rate, ERA5 underestimates

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

the total accumulation (334 to 441 mm). When matching phase precisely (EO_{phase}), ERA5 does slightly better in producing the observed accumulated precipitation.

Finally, we look at the most constrained hours when all three instruments observe precipitation (IEO_{precip}) (row 8) and concurrent phase (IEO_{phase}) (row 9). About 12% of all records are consistent in total precipitation and less than 9% have a consistent phase among all three products. During these times (phase consistency), IMERG records twice as much total precipitation as the ODM470 and 10 times more snow/ice. Mixed phase accumulated precipitation in ERA5 tends to be overestimated by up to 20% during IEO_{phase} periods compared to the disdrometer. The accumulated total precipitation during the IEO_{phase} period accounts for the majority (50%) of the accumulated amounts from the three individual datasets (rows 1-3), suggesting that IEO_{phase} periods may have captured the heavy and widespread precipitation events where key precipitation features are more consistently represented in all products.

Given the relatively poor agreement between the IMERG and ODM470 records, it is worthwhile to examine the contribution of various inputs into the IMERG algorithm at the concurrent times (IO_{precip}) (See the supplementary material for a further explanation). The IMERG product was never derived from the IR-only technique. Rather the IMERG algorithm more commonly employed the microwave (PWM-only) product (44% only) and the morphing-only product (29%). About 27% of the concurrent hours used the IR technique in combination with other records (mixture), commonly with only a small weight; 75% of these records had a weight below 7%. Sources of error in the concurrent ERA5 precipitation (EO_{precip}) are more difficult to untangle with 64% of the ERA5 precipitation being produced by the microphysics and 36% from the convective parameterization. Similar results are found when analyzing IMERG and ERA5 precipitation for all collocated datasets.

A probability distribution function (PDF) of the precipitation rate was calculated for all collocated datasets (Figure 7). Given the nature of each dataset, we cannot expect a complete match of the PDFs even in the best-case scenario, but any systematic bias will be identifiable. The PDFs suggest that, compared to the ODM470, ERA5 tends to overestimate the FoO of precipitation events with intensities below 0.1 mm hr^{-1}, whereas IMERG tends to overestimate the frequency for intensities above 0.5 mm hr^{-1} (Figure 7a). We further explored the IMERG precipitation rate above 0.5 mm hr^{-1} tracking the input back to mainly PMW-only algorithm.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

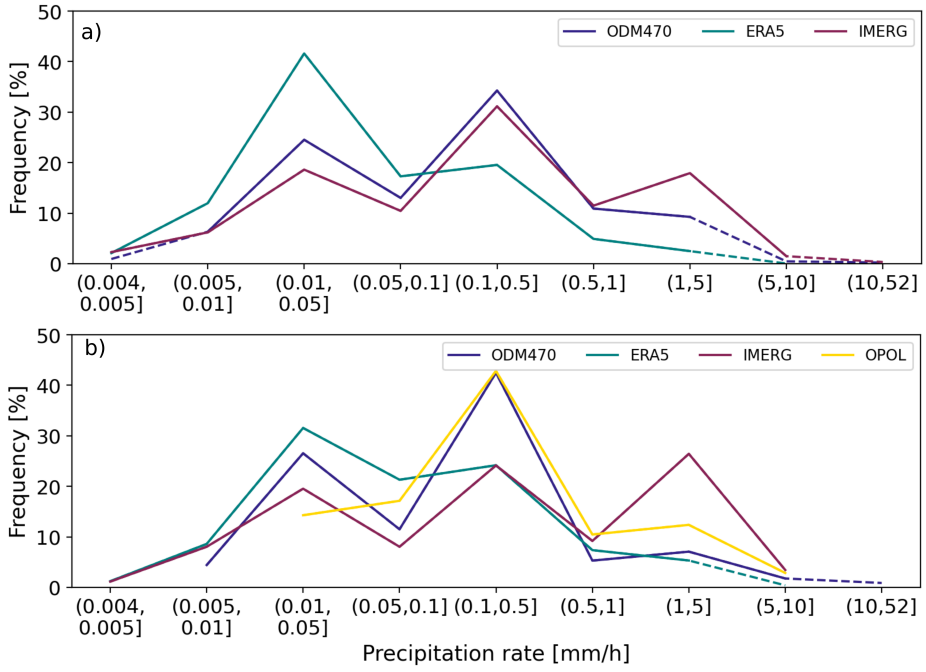


FIG. 7. Probability distribution function (PDF) of precipitation for IMERG (brown), ERA5 (green), and ODM470 (purple) during six voyages in the SO between January 2016 and March 2018. a) all thermodynamic phases. b) same as a) but only for liquid events detected by the disdrometer while OceanPOL (yellow) was functioning. Dashed lines show frequencies below 1%.

The ODM470 or the OceanPOL detected about 305 hours of rain events throughout the 2018 voyages, the OceanPOL indicates a PDF similar to the ODM470, especially for precipitation in the range between 0.05 and 1 mm hr^{-1} (Figure 7b). Differences outside this range can be associated with the sampling differences, instrument uncertainties in detecting drizzle/light precipitation, and/or possibly the radar retrieval algorithm calibrated for the tropical conditions. Overall, the statistical analysis and the examples presented in section 3 suggest that the OceanPOL and the ODM470 precipitation rate estimates are in reasonable agreement, which justifies the use of the OceanPOL dataset for a spatial analysis presented in Section 5.

In summary, the quantitative estimates of snow and mixed-phase precipitation from IMERG and ERA5 are subject to notable errors with differing characteristics. IMERG detects only 20% of the snow and mixed phase hours compared to the ODM470, while ERA5 detected up to 50% of the hours. For accumulated liquid precipitation, ERA5 seems to agree with the disdrometer

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

records better than IMERG. However, this is largely a result of compensating errors where ERA5 overestimates the frequency and underestimates the intensity of the precipitation. IMERG, on the other hand, tends to overestimate precipitation intensity, which results in a higher accumulated precipitation amount.

5. Spatial analysis: OceanPOL

Bearing in mind the inherent limitations of the OceanPOL dataset as discussed in Section 2, the spatial coverage of the C-band radar eliminates the "point-to-area" problem noted earlier, and the realism of the rain rate retrievals (e.g. Figure 7) allows the use of this dataset for a spatial verification of the ERA5 and IMERG products. Here we use a total of 305 liquid precipitation hours collected with the OceanPOL in 2018 to calculate the Fractional Skill Score (FSS) against the ERA5 and IMERG data. For consistency, we have only included precipitation values above the trace threshold (0.1 mm day^{-1}) at 60 min and 0.25^o resolution for this analysis. During the selected period, ERA5 produces a maximum rain rate of 6 mm hr^{-1}, which is much lower than the maximum rain rate from the IMERG (15 mm hr^{-1}) and the OceanPOL (18 mm hr^{-1}) records. Rain rates above 6 mm hr^{-1} are present 3% (1%) of the time in IMERG (OceanPOL). ERA5 produces rain rate values above 1 mm hr^{-1} less than 5% of the time, while these values are present 33% and 25% of the time in the IMERG and OceanPOL datasets (not shown). Given the prevalence of low rain intensities, we calculate the FSS for precipitation thresholds only up to 6 mm hr^{-1}. The spatial scale considered in our analysis is also constrained by the spatial range of the OceanPOL observations, which is between 28 and 194 km.

As expected, the FSS results show better skills at larger spatial scales and lower rain rates, except for the very light rain which may not be detectable by the C-band radar, although this does not guarantee that ERA5 and IMERG are correct for these low rain rates (Figure 8). ERA5 appears to have more skill for smaller spatial scales than IMERG (55km vs 83km). It is worth noting that the skill of ERA5 increases with the spatial scale (i.e., useful scores are produced for a wider range of precipitation thresholds as the spatial scale increases). The above-mentioned tendency is less consistent for IMERG. If we use 0.5 as the useful FSS limit (including light red colors in Figure 8) IMERG appears to be largely insensitive to the spatial scale for ranges between 55 and 194 km and for thresholds below 1.4 mm hr^{-1}, while ERA5 skillful threshold increases to 3.5 mm hr^{-1}.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

By increasing the precipitation threshold and thereby limiting the skill analysis to more intense precipitation events, our results suggest that ERA5 and IMERG are no longer skillful at any spatial scale for precipitation thresholds above 2.7 mm hr^{-1}.

In general, ERA5 has better FSS scores than IMERG at large spatial scales for precipitation thresholds between 0.07 and 0.6 mm hr^{-1}, which is not particularly surprising given the high frequency of drizzling events present in ERA5. As noted earlier, the overestimation of precipitation intensities in IMERG does not necessarily translate into a better skill in detecting strong precipitation events.

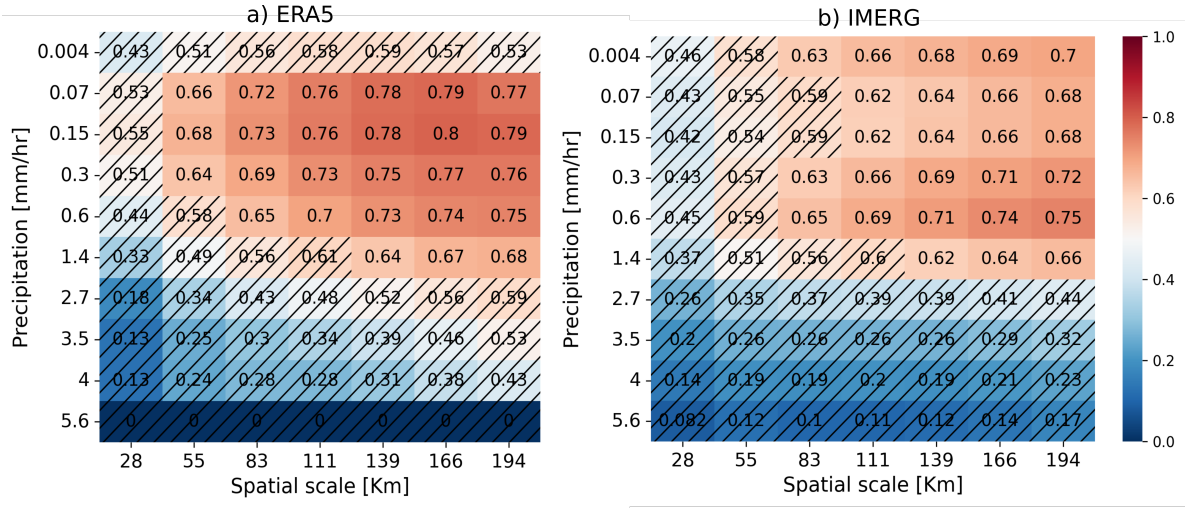


FIG. 8. ERA5 and IMERG Fraction Skill Score (FSS) calculated using the OceanPOL precipitation estimates. (Left) OceanPOL vs. ERA5 and (right) OceanPOL vs. IMERG. All datasets have been averaged to a 60 min, 0.25°x0.25° resolution. Hashed areas indicate scores below FSS_{useful}.

6. Precipitation evaluation by synoptic conditions

It is reported in Montoya Duque et al. (2022) that distinct cloud and precipitation regimes are found to correspond to the seven synoptic clusters over the SO as determined by the k-means clustering from Truong et al. (2020). Therefore, we extend our analysis to examine the skill of ERA5 and IMERG precipitation products under these different weather regimes for the 4346 collocated hours (Table 3). The OceanPOL is excluded in this section as at the time of writing not enough data is available from the OceanPOL to allow for a statistically robust analysis.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

TABLE 3. ERA5, IMERG (1 hour,0.25°) and ODM470 disdrometer collocated values as in rows 1-3 in Table 2. Data are differentiated by seven synoptic conditions: W1 (warm air advection in pre-frontal areas north of the polar front), M1 (high pressure), M2 (cold front), M3 (post-frontal), M4 (warm front), C1 (high-latitude cyclone center), and C2 (Coastal Antarctica). Frequency of Ocurrence (FoO, %), quartiles 25, 50 and 75 (Q_{25}, Q_{50}, and Q_{75}, [mm/hr]) and accumulated precipitation (Acumm [mm]). Hours include non-precipitating periods.

	Hours	FoO [%]			Q_{25}, Q_{50}, Q_{75} [mm/hr]			Accum [mm]		
Synoptic	Total	ERA5	IMERG	ODM470	ERA5	IMERG	ODM470	ERA5	IMERG	ODM470
Warm advection-W1	1150	57	19	19	0.01, 0.03, 0.18	0.02, 0.10, 0.48	0.03, 0.07, 0.18	80	136	69
High Pressure-M1	495	24	4	2	0.01, 0.02, 0.03	0.03, 0.06, 0.21	0.03, 0.05, 0.08	3	2	0.6
Cold Front-M2	630	87	25	40	0.02, 0.05, 0.13	0.03, 0.15, 0.52	0.04, 0.14, 0.39	78	58	72
Post-frontal-M3	492	76	13	25	0.02, 0.04, 0.07	0.03, 0.06, 0.14	0.02, 0.06, 0.17	25	14	24
Warm Front-M4	285	96	66	73	0.10, 0.28, 0.79	0.21, 0.66, 1.80	0.11, 0.34, 0.86	147	258	138
Cyclone center-C1	1157	70	27	27	0.02, 0.04, 0.12	0.07, 0.20, 0.60	0.04, 0.11, 0.33	103	143	145
Antarctic Coast-C2	137	12	7	0	0.01, 0.01, 0.02	0.07, 0.11, 0.19	0	0.2	2	0

Looking first at the IMERG dataset, the precipitation frequency is about half of what was detected by the disdrometer under the cold front (M2) and post-frontal conditions (M3), although the quartiles of precipitation intensity are more comparable (Table 3). Under the warm air advection in pre-frontal areas (W1), IMERG detects a precipitation frequency similar to the disdrometer record, but the upper quartile of precipitation intensity is overestimated, resulting in an overestimate of the accumulated amount (about 200% more). Overall, warm frontal (M4) precipitation intensity is overestimated by about a factor of two, but the frequency is underestimated. Precipitation frequency associated with the high-latitude cyclone (C1) is similar to the disdrometer record, but the intensity quartiles are overestimated. Unlike the warm frontal precipitation, these conditions do not translate to an overestimated accumulated precipitation. IMERG detects a small amount of precipitation under the Antarctic Coastal condition (C2) although no precipitation was recorded by the disdrometer. We can also note that Antarctic Coast (C2) and the high-latitude cyclone center (C1) precipitation are estimated either from the morphing-only or the PMW-only method (see Table S1) as the IR scheme is rarely used when high oblique viewing angles are present (Tan et al. 2019). Warm frontal precipitation, where IMERG is prone to marked precipitation intensity overestimation, comes mainly from the PMW-only method, consistent with the bias found in the high-latitude case study (section 3b). High-pressure (M1) and warm air advection in pre-frontal areas (W1) precipitation estimates are mostly from a mixture of techniques inluding IR to some

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

extent. Post-frontal (M3) precipitation is mostly from PMW-only or morphing-only techniques and cold frontal (M2) precipitation can come from any of the three analyzed sources.

ERA5 precipitation, on the other hand, tends to overestimate the frequency of precipitation events compared to the disdrometer for all seven clusters. However, the overestimation of frequency tends to almost compensate the underestimation of precipitation intensity, resulting in similar accumulated amounts. High-pressure conditions (M1) are associated with the greatest underestimation of intensity and overestimation of frequency. ERA5 also tends to record the lowest intensity events under the Antarctic coastal conditions, while the disdrometer indicates dry conditions. More in depth analysis from ERA5 would be needed to determine the source of errors.

In summary, our analysis suggests that IMERG tends to overestimate the higher intensity (upper quartile) precipitation values for all clusters except the post-frontal condition (M3). However, during M3 conditions the precipitation frequency is underestimated. During warm frontal (M4) and high-latitude cyclone (C1) conditions, IMERG precipitation intensity overestimation is prevalent. ERA5 overestimates precipitation frequency for all clusters, which leads to an overestimation of accumulated precipitation even when the intensity is typically lower than the shipborne estimates.

7. Discussion and concluding remarks

Quantitative precipitation estimates over the Southern Ocean (SO) continue to suffer from a high degree of uncertainty with large differences present amongst both satellite-based and reanalysis products. This uncertainty arises from the absence of long-term, high-quality observations of precipitation suitable for evaluation across a range of temporal and spatial scales. Recent voyages on the Australian RV Investigator, however, have provided new observations of precipitation made with both the OceanRAIN maritime disdrometer (ODM470) and a dual-polarization C-band weather radar (OceanPOL).

In this study, we employed these unprecedented observations to evaluate the Global Precipitation Mission (GPM) Integrated Multi-satellitE Retrievals (IMERG) and the ECMWF reanalysis (ERA5) precipitation products at a scale of 60 min and 0.25°, i.e., at the mesoscale process level. A summary of key findings is as below:

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

- ERA5 generally overestimates the precipitation frequency compared to the OceanRAIN disdrometer, whereas precipitation intensity is underestimated. This behavior is exacerbated under high-pressure conditions.
- IMERG tends to overestimate the precipitation intensity compared to the OceanRAIN disdrometer, especially for warm frontal and high-latitude cyclone precipitation. Under the post-frontal conditions, IMERG underestimates the precipitation frequency while showing similar precipitation intensities.
- ERA5 has better skill in detecting precipitation events of different spatial scales than IMERG when compared against the OceanPOL C-band weather radar.
- There is little agreement in the phase classification between the OceanRAIN disdrometer, IMERG, and ERA5. The comparisons are complicated by the various assumptions for phase classification in the different datasets.

Previous evaluation of various precipitation products over the SO has largely been limited to MAC Island. Using a newly developed blended observational dataset, Tansey et al. (2022) found that the annual precipitation frequency is 44% ± 4% at MAC Island and 27% ± 3% during Austral summer. Our analysis across the broader extent of the SO suggests comparable statistics, although the available datasets we have employed are somewhat biased towards the warm seasons. More campaign observations spanning across different seasons would be desirable to more fully account for potential seasonal cycle in the precipitation features (Lang et al. 2020; Tansey et al. 2022).

Our analysis suggests that ERA5 commonly overestimates the frequency of precipitation and underestimates the intensity of precipitation, although the accumulated precipitation compares reasonably well with the OceanRAIN disdrometer data, with an underestimation of only approximately 2.3%. This is consistent with recent findings in other oceanic regions (Behrangi et al. 2016; Wang et al. 2018; Burdanowitz et al. 2019; Naud et al. 2020). These high-frequency and low-intensity events are commonly referred to as the ”drizzle problem”, which are frequently found in Global Circulation Models and often linked to deficiencies in the parameterized convection and/or microphysics. While ERA5 has been suggested to be generally superior to its predecessor ERA-Interim, our analysis suggests that the long-standing drizzle errors remain prevalent over the SO. These biases have further implications for the modeled cloud cover and latent heat flux,

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

which have a profound influence on the energy budget and atmospheric circulation (Terai et al. 2016; Barrett et al. 2020), particularly over the SO where the model bias in the energy budget is known to be pronounced (Bodas-Salcedo et al. 2014). Encouragingly, our analysis suggests that the ERA5 precipitation intensity estimates are more accurate during heavy and widespread precipitation events over the SO, although these events are less common.

Although previous studies have suggested that IMERG tends to under-represent the global mean oceanic precipitation as estimated from OceanRAIN (Khan and Maggioni 2019), our analysis shows that IMERG tends to overestimate precipitation intensity over the SO, most notably under the organized synoptic-scale systems such as warm fronts and high-latitude cyclones. This is consistent with some analysis of mid-latitude precipitation using rain gauge data over land areas (Xu et al. 2017; Moazami and Najafi 2021) and in moist environments over extratropical oceans (Naud et al. 2018, 2020). Our study found that IMERG underestimates precipitation frequency and intensity under post-frontal conditions (M3), where shallow convective clouds prevail (Lang et al. 2020, 2021; Montoya Duque et al. 2022). This suggests that the retrieval errors in the IMERG product are likely regime-dependent. Also, as suggested by Naud et al. (2020), we find that IMERG has greater challenges in properly detecting frozen precipitation. Our analysis suggests that the IMERG precipitation rate estimates over the SO are more heavily influenced by the PMW-only and the morphing-only schemes, and that the overestimate of precipitation rate is more strongly tied to the PMW-only and the 'mixture' methods. While PMW precipitation estimates have been reported to be less accurate over frozen surfaces and in detecting light precipitation (Tan et al. 2019; Naud et al. 2020), tracing back the exact sources of these errors over the SO is out of the scope of this study, especially considering the 10 PMW satellites and different processing techniques involved (Huffman et al. 2019a,b). We hypothesize that the heavy reliance on the PMW-only technique may help explain the bias of this product under the warm front and high-latitude cyclone conditions over this region, where the advection-driven, multi-layer clouds are commonly present between 6-8 km (Mace et al. 2009; Truong et al. 2020; Montoya Duque et al. 2022). These multi-layer clouds, which are not necessarily heavily precipitating, might produce a bulk radiative signature (e.g. brightness temperature) that is similar to the deep convective systems. The morphing scheme is limited by the temporal and spatial resolution since it is based on the MERRA-2 reanalysis (1 h, 2.5°) for computing the motion vectors, which are then interpolated to 30 min and 0.1° (Tan

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

et al. 2019)). Also, the morphing method faces challenges in representing short-lived or rapidly moving precipitation systems (Turk et al. 2008). Further, cloud and precipitation microphysics assumptions widely used in the satellite/model algorithms, such as particle size distributions, also need to be revisited for the SO environment, as suggested by Protat et al. (2019a,b); Li et al. (2022).

Both ERA5 and IMERG poorly represent the dry surface conditions near the Antarctic Coast (cluster C2), a unique area characterized by dry katabatic winds, which often favors strong evaporation/sublimation of precipitation near the surface (Truong et al. 2020). Despite other factors, the limited resolutions of these products make it difficult to represent the time-space variability of precipitation associated with the complex coastal and orographic processes. Nevertheless, we note that surface precipitation measurements in this environment are also limited. On the other hand, ERA5 and IMERG can perform differently at lower spatial and temporal resolutions as well as in other seasons (Tan et al. 2017; Tansey et al. 2022). We did not include those analysis since in-situ data is limited in this region.

In line with findings in Manton et al. (2020), our analysis suggests that ERA5 quantitative precipitation estimates are overall more accurate than the IMERG product over the SO. The spatial precipitation patterns are also better captured by ERA5 when evaluated against the (limited) OceanPOL radar data. However, as noted earlier, the scarcity of precipitation measurements over the SO, particularly at the high latitudes, remains a significant impediment to the continuous effort of evaluating and improving the precipitation products in this remote region. Enhancing the long-term supports for surface measurement capabilities, such as OceanRAIN and OceanPOL, will continue to be the key to improving our confidence in understanding precipitation as well as a wide range of climate processes.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

Acknowledgments. The Australian Research Council discovery grant DP190101362 supported this work. The ARC Centre of Excellence for Climate Extremes also supported Y. Huang (CE170100023). The Securing Antarctica's Environmental Future (SAEF) program supported S.T. Siems. The authors acknowledge the thorough work of the Bureau of Meteorology (BoM) and CSIRO teams to collect, post-process, and make available the datasets collected in different voyages with the RV Investigator, as well as the NASA and ECMWF teams for producing the IMERG and ERA5 datasets. We especially thank Dr. Alain Protat and Dr. Valentin Louf from the BoM for their great efforts in quality control, producing the precipitation estimates from the OceanPOL radar, and providing the Himawari 8 dataset. Also, we thank the National Computational Infrastructure (NCI) for providing/storing the OceanPOL, IMERG, ERA5, and Himawari 8 data, and Dr. Son Truong for providing the k-means cluster center values for the synoptic classification. Lastly, we thank the reviewers for their insightful comments which helped us improving this paper.

Data availability statement. The OceanRAIN disdrometer data used in the study are available at CSIRO via `https://doi.org/10.25919/5f688fcc97166` with creative commons attribution 4.0 International. The OceanPOL V2020 weather radar is available on the National Computing Infrastructure (NCI) via `https://dx.doi.org/10.25914/5fc4975c7dda8`. The CAPRICORN II data used in the case studies are available at CSIRO via `https://doi.org/10.25919/5f688fcc97166` with creative commons attribution 4.0 International. ERA5 hourly data (10.24381/cds.bd0915c6) and the IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degrees V06 (10.5067/GPM/IMERG/3B-HH/06) datasets were provided by NCI. The NCI downloads and stores ERA5 and IMERG data from ECMWF and NASA/Goddard official downloading websites and makes them available through user registration. The Himawari-8 satellite product was provided by Dr. Alain Protat upon request; the dataset was processed to include high-latitude information during the CAPRICORN field campaign by the Australian Bureau of Meteorology (BoM). All used datasets are available in the in-text data citation references: Huffman et al. (2021), Hersbach et al. (2018),Protat (2018), UCAR/NCAR EOL (2018), Bessho et al. (2016),Louf and Protat (2020) and Protat and CSIRO (2020).

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

References

Barrett, A. P., J. C. Stroeve, and M. C. Serreze, 2020: Arctic ocean precipitation from atmospheric reanalyses and comparisons with north pole drifting station records. *Journal of Geophysical Research: Oceans*, **125**, https://doi.org/10.1029/2019JC015415.

Behrangi, A., and Y. Song, 2020: A new estimate for oceanic precipitation amount and distribution using complementary precipitation observations from space and comparison with gpcp. *Environmental Research Letters*, **15**, https://doi.org/10.1088/1748-9326/abc6d1.

Behrangi, A., Y. Tian, B. H. Lambrigtsen, and G. L. Stephens, 2014: What does CloudSat reveal about global land precipitation detection by other spaceborne sensors? *Water Resources Research*, **50**, https://doi.org/10.1175/JCLI-D-13-00169.1.

Behrangi, A., and Coauthors, 2016: Status of high-latitude precipitation estimates from observations and reanalyses. *Journal of Geophysical Research: Atmospheres*, **121**, https://doi.org/10.1002/2015JD024546.

Berry, G., M. J. Reeder, and C. Jakob, 2011: A global climatology of atmospheric fronts. *Geophysical Research Letters*, **38**, https://doi.org/10.1029/2010GL046451.

Bessho, K., and Coauthors, 2016: An introduction to himawari-8/9 japan's new-generation geostationary meteorological satellites. *Journal of the Meteorological Society of Japan*, **94**, https://doi.org/10.2151/jmsj.2016-009.

Bodas-Salcedo, A., and Coauthors, 2014: Origins of the Solar Radiation Biases over the Southern Ocean in CFMIP2 Models. *Journal of Climate*, **27**, https://doi.org/10.1175/JCLI-D-13-00169.1.

Boisvert, L. N., M. A. Webster, A. A. Petty, T. Markus, R. I. Cullather, and D. H. Bromwich, 2020: Intercomparison of precipitation estimates over the southern ocean from atmospheric reanalyses. *Journal of Climate*, **33**, https://doi.org/10.1175/JCLI-D-20-0044.1.

Bumke, K., K. Fennig, A. Strehz, R. Mecking, and M. Schröder, 2012: Hoaps precipitation validation with ship-borne rain gauge measurements over the baltic sea. *Tellus A: Dynamic Meteorology and Oceanography*, **64**, https://doi.org/10.3402/tellusa.v64i0.18486.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

Bumke, K., G. König-Langlo, J. Kinzel, and M. Schröder, 2016: Hoaps and era-interim precipitation over the sea: validation against shipboard in situ measurements. *Atmospheric Measurement Techniques*, **9**, https://doi.org/10.5194/amt-9-2409-2016.

Burdanowitz, J., S. A. Buehler, S. Bakan, and C. Klepp, 2019: The sensitivity of oceanic precipitation to sea surface temperature. *Atmospheric Chemistry and Physics*, **19**, https://doi.org/10.5194/acp-19-9241-2019.

Ceppi, P., M. D. Zelinka, and D. L. Hartmann, 2016: The response of the southern hemispheric eddy-driven jet to future changes in shortwave radiation in cmip5. *Geophysical Research Letters*, **41**, https://doi.org/10.1002/2014GL060043.

CSIRO, 2022: RV Investigator. CSIRO, accessed: 2022-02-02, https://mnf.csiro.au/en/RV-Investigator.

Delanoë, J., and Coauthors, 2016: BASTA: A 95-GHz FMCW Doppler Radar for Cloud and Fog Studies. *Journal of Atmospheric and Oceanic Technology*, **33**, https://doi.org/10.1175/JTECHD-15-0104.1.

Gettelman, A., and Coauthors, 2019: High climate sensitivity in the community earth system model version 2 (cesm2). *Geophysical Research Letters*, **46**, https://doi.org/10.1029/2019GL083978.

Hersbach, H., and Coauthors, 2018: ERA5 hourly data on single levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), accessed: 2021-06-01, https://doi.org/10.24381/cds.adbb2d47.

Hersbach, H., and Coauthors, 2020: The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**, https://doi.org/10.1002/qj.3803.

Huang, Y., A. Protat, S. T. Siems, and M. J. Manton, 2015: A-Train Observations of Maritime Midlatitude Storm-Track Cloud Systems: Comparing the Southern Ocean against the North Atlantic. *Journal of Climate*, **28**, https://doi.org/10.1175/JCLI-D-14-00169.1.

Huffman, G., E. Stocker, D. Bolvin, E. Nelkin, and J. Tan, 2021: GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06 Dataset. NASA, accessed: 2021-01-06, https://gpm.nasa.gov/data/directory, https://doi.org/10.5067/GPM/IMERG/3B-HH/06.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

Huffman, G. J., D. T. Bolvin, E. J. Nelkin, and J. Tan, 2019a: Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. Tech. rep., Global Precipitation Measurement - NASA.

Huffman, G. J., and Coauthors, 2019b: Algorithm Theoretical Basis Document (ATBD) Version 06. Tech. rep., Global Precipitation Measurement - NASA.

Keil, C., and G. C. Craig, 2009: A displacement and amplitude score employing an optical flow technique. *Weather and Forecasting*, **24**, https://doi.org/10.1175/2009WAF2222247.1.

Khan, S., and V. Maggioni, 2019: Assessment of level-3 gridded global precipitation mission (gpm) products over oceans. *Remote Sensing*, **11**, https://doi.org/10.3390/rs11030255.

Klepp, C., 2020: OceanRAIN Release 2.0 User Manual. Where Geosciences Meets Art, accessed: 2022-11-01.

Klepp, C., K. Bumke, S. Bakan, and P. Bauer, 2010: Ground validation of oceanic snowfall detection in satellite climatologies during lofzy. *Tellus A: Dynamic Meteorology and Oceanography*, **62**, https://doi.org/10.1111/j.1600-0870.2009.00459.x.

Klepp, C., P. A. Kucera, J. Burdanowitz, and A. Protat, 2020: *OceanRAIN The Global Ocean Surface-Reference Dataset for Characterization, Validation and Evaluation of the Water Cycle*, 655–674. Springer International Publishing, https://doi.org/10.1007/978-3-030-35798-6_10.

Klepp, C., and Coauthors, 2018: OceanRAIN, a new in-situ shipboard global ocean surface-reference dataset of all water cycle components. *Scientific Data*, **5**, https://doi.org/10.1038/sdata.2018.122.

Lang, F., Y. Huang, A. Protat, S. C. H. Truong, S. T. Siems, and M. J. Manton, 2021: Shallow convection and precipitation over the southern ocean: A case study during the capricorn 2016 field campaign. *Journal of Geophysical Research: Atmospheres*, **126**, https://doi.org/10.1029/2020JD034088.

Lang, F., Y. Huang, S. Siems, and M. Manton, 2018: Characteristics of the marine atmospheric boundary layer over the Southern Ocean in response to the synoptic forcing. *Journal of Geophysical Research: Atmospheres*, **123**, https://doi.org/10.1029/2018JD028700.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

Lang, F., Y. Huang, S. T. Siems, and M. J. Manton, 2020: Evidence of a diurnal cycle in precipitation over the southern ocean as observed at macquarie island. *Atmosphere*, **11**, https://doi.org/10.3390/atmos11020181.

Lavers, D. A., A. Simmons, F. Vamborg, and M. J. Rodwell, 2022: An evaluation of era5 precipitation for climate monitoring. *Quarterly Journal of the Royal Meteorological Society*, **148**, https://doi.org/10.1002/qj.4351.

Li, Z., G. Tang, P. Kirstetter, S. Gao, J.-L. Li, Y. Wen, and Y. Hong, 2022: Evaluation of gpm imerg and its constellations in extreme events over the conterminous united states. *Journal of Hydrology*, **606**, https://doi.org/https://doi.org/10.1016/j.jhydrol.2021.127357.

Loew, A., and Coauthors, 2017: Validation practices for satellite-based earth observation data across communities. *Reviews of Geophysics*, **55**, https://doi.org/10.1002/2017RG000562.

Louf, V., and A. Protat, 2020: OceanPOL Weather Radar Dataset. V1. National Computing Infrastructure, accessed: 2021-10-03, https://doi.org/10.25914/5fc4975c7dda8.

Louf, V., A. Protat, R. A. Warren, S. M. C., D. B. Wolff, S. Raunyiar, C. Jakob, and W. A. Petersen, 2019: An integrated approach to weather radar calibration and monitoring using ground clutter and satellite comparisons. *Journal of Atmospheric and Oceanic Technology*, **36**, https://doi.org/10.1175/JTECH-D-18-0007.1.

Mace, G. G., A. Protat, S. Benson, and P. McGlynn, 2022: Inferring the properties of snow in southern ocean shallow convection and frontal systems using dual polarization c-band radar. *Journal of Applied Meteorology and Climatology*, https://doi.org/10.1175/JAMC-D-22-0097.1.

Mace, G. G., Q. Zhang, M. Vaughan, R. Marchand, G. Stephens, C. Trepte, and D. Winker, 2009: A description of hydrometeor layer occurrence statistics derived from the first year of merged cloudsat and calipso data. *Journal of Geophysical Research: Atmospheres*, **114**, https://doi.org/10.1029/2007JD009755.

Manton, M. J., Y. Huang, and S. T. Siems, 2020: Variations in precipitation across the southern ocean. *Journal of Climate*, **33**, https://doi.org/10.1175/JCLI-D-20-0120.1.

McFarquhar, G. M., and Coauthors, 2021: Observations of Clouds, Aerosols, Precipitation, and Surface Radiation over the Southern Ocean: An Overview of CAPRICORN, MARCUS, MICRE,

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

and SOCRATES. *Bulletin of the American Meteorological Society*, **102**, https://doi.org/10.1175/BAMS-D-20-0132.1.

Mittermaier, M., and N. Roberts, 2010: Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score. *Weather and Forecasting*, **25**, https://doi.org/10.1175/2009WAF2222260.1.

Moazami, S., and M. Najafi, 2021: A comprehensive evaluation of gpm-imerg v06 and mrms with hourly ground-based precipitation observations across canada. *Journal of Hydrology*, **594**, https://doi.org/10.1016/j.jhydrol.2020.125929.

Montoya Duque, E., Y. Huang, S. T. Siems, P. T. May, A. Protat, and G. M. McFarquhar, 2022: A characterization of clouds and precipitation over the southern ocean from synoptic to micro scales during the capricorn field campaigns. *Journal of Geophysical Research: Atmospheres*, **127**, https://doi.org/10.1029/2022JD036796.

Murray, R. J., and I. Simmonds, 1991: A numerical scheme for tracking cyclone centres from digital data. part i: development and operation of the scheme. *Australian Meteorological Magazine*, **39**.

Naud, C. M., J. F. Booth, M. Lebsock, and M. Grecu, 2018: Observational constraint for precipitation in extratropical cyclones: Sensitivity to data sources. *Journal of Applied Meteorology and Climatology*, **57**, https://doi.org/JAMC-D-17-0289.1.

Naud, C. M., J. Jeyaratnam, J. F. Booth, M. Zhao, and A. Gettelman, 2020: Evaluation of modeled precipitation in oceanic extratropical cyclones using imerg. *Journal of Climate*, **33**, https://doi.org/10.1175/JCLI-D-19-0369.1.

Protat, A., 2018: Capricorn r/v investigator 95-ghz doppler cloud radar (basta) level 2 data. Accessed: 2020-12-01, https://doi.org/10.26023/FHQN-F907-NT0T.

Protat, A., and CSIRO, 2020: RV Investigator BOM Atmospheric Data Overview (2016 onwards). v2. Marine National Facility, accessed: 2021-06-01, https://doi.org/10.25919/5f688fcc97166.

Protat, A., C. Klepp, V. Louf, W. A. Petersen, S. P. Alexander, A. Barros, J. Leinonen, and G. G. Mace, 2019a: The latitudinal variability of oceanic rainfall properties and its implication

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

for satellite retrievals: 1. drop size distribution properties. *Journal of Geophysical Research: Atmospheres*, **124**, https://doi.org/10.1029/2019JD031010.

Protat, A., C. Klepp, V. Louf, W. A. Petersen, S. P. Alexander, A. Barros, J. Leinonen, and G. G. Mace, 2019b: The latitudinal variability of oceanic rainfall properties and its implication for satellite retrievals: 2. the relationships between radar observables and drop size distribution parameters. *Journal of Geophysical Research: Atmospheres*, **124**, https://doi.org/10.1029/2019JD031011.

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, **136**, https://doi.org/10.1175/2007MWR2123.1.

Siems, S., Y. Huang, and M. Manton, 2022: Southern ocean precipitation: Toward a process level understanding. *WIREs Climate Change*, **e800**, https://doi.org/10.1002/wcc.800.

Skofronick-Jackson, G., and Coauthors, 2017: The global precipitation measurement (gpm) mission for science and society. *Bulletin of the American Meteorological Society*, **98**, https://doi.org/10.1175/BAMS-D-15-00306.1.

Tan, J., G. J. Huffman, D. T. Bolvin, and E. J. Nelkin, 2019: Imerg v06: Changes to the morphing algorithm. *Journal of Atmospheric and Oceanic Technology*, **36**, https://doi.org/https://doi.org/10.1175/JTECH-D-19-0114.1.

Tan, J., W. A. Petersen, P.-E. Kirstetter, and Y. Tian, 2017: Performance of imerg as a function of spatiotemporal scale. *Journal of Hydrometeorology*, **18**, https://doi.org/10.1175/JHM-D-16-0174.1.

Tansey, E., R. Marchand, A. Protat, S. P. Alexander, and S. Ding, 2022: Southern ocean precipitation characteristics observed from cloudsat and ground instrumentation during the macquarie island cloud & radiation experiment (micre): April 2016 to march 2017. *Journal of Geophysical Research: Atmospheres*, **127**, https://doi.org/10.1029/2021JD035370.

Terai, C., P. Caldwell, and S. Klein, 2016: Why do climate models drizzle too much and what impact does this have. *AGU Fall Meeting Abstracts*, Vol. 2016.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

Thompson, E. J., S. A. Rutledge, B. Dolan, V. Chandrasekar, and B. L. Cheong, 2014: A dual-polarization radar hydrometeor classification algorithm for winter precipitation. *Journal of Atmospheric and Oceanic Technology*, **31**, https://doi.org/10.1175/JTECH-D-13-00119.1.

Thompson, E. J., S. A. Rutledge, B. Dolan, and M. Thurai, 2015: Drop size distributions and radar observations of convective and stratiform rain over the equatorial indian and west pacific oceans. *Journal of the Atmospheric Sciences*, **72**, https://doi.org/10.1175/JAS-D-14-0206.1.

Thompson, E. J., S. A. Rutledge, B. Dolan, M. Thurai, and V. Chandrasekar, 2018: Dual-polarization radar rainfall estimation over tropical oceans. *Journal of Applied Meteorology and Climatology*, **57**, https://doi.org/10.1175/JAMC-D-17-0160.1.

Truong, S., Y. Huang, F. Land, M. Messmer, I. Simmonds, S. Siems, and M. Manton, 2020: A climatology of the marine atmospheric boundary layer over the Southern Ocean from four field campaigns during 2016 – 2018. *Journal of Geophysical Research Atmospheres*, **125**, https://doi.org/10.1029/2020JD033214.

Truong, S. C. H., Y. Huang, S. T. Siems, M. J. Manton, and F. Lang, 2022: Biases in the thermodynamic structure over the southern ocean in era5 and their radiative implications. *International Journal of Climatology*, **n/a**, https://doi.org/10.1002/joc.7672.

Turk, F. J., P. Arkin, E. E. Ebert, and M. R. P. Sapiano, 2008: Evaluating high-resolution precipitation products. *Bulletin of the American Meteorological Society*, **89**.

UCAR/NCAR EOL, 2018: Ncar/eol iss radiosonde data. version 1.0. Accessed: 2020-12-01, https://doi.org/10.5065/D69P30HG.

Wang, C., G. Tang, Z. Han, X. Guo, and Y. Hong, 2018: Global intercomparison and regional evaluation of gpm imerg version-03, version-04 and its latest version-05 precipitation products: Similarity, difference and improvements. *Journal of Hydrology*, **564**, https://doi.org/10.1016/j.jhydrol.2018.06.064.

Wang, Z., S. T. Siems, D. Belusic, M. J. Manton, and Y. Huang, 2015: A climatology of the precipitation over the southern ocean as observed at macquarie island. *Journal of Applied Meteorology and Climatology*, **54**, https://doi.org/10.1175/JAMC-D-14-0211.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.

Xu, R., F. Tian, L. Yang, H. Hu, H. Lu, and A. Hou, 2017: Ground validation of gpm imerg and trmm 3b42v7 rainfall products over southern tibetan plateau based on a high-density rain gauge network. *Journal of Geophysical Research: Atmospheres*, **122**, https://doi.org/10.1002/2016JD025418.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-23-0039.1.