SPECIAL
Process-Oriented Model Diagnostics
COLLECTION

Evaluation of Modeled Precipitation in Oceanic Extratropical Cyclones Using IMERG🖉

CATHERINE M. NAUD

Applied Physics and Applied Mathematics, Columbia University/NASA GISS, New York, New York

JEYAVINOTH JEYARATNAM AND JAMES F. BOOTH

Earth and Atmospheric Sciences, City University of New York, City College and the Graduate Center, New York, New York

MING ZHAO

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

ANDREW GETTELMAN

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 23 May 2019, in final form 13 August 2019)

ABSTRACT

Using a high-spatial- and high-temporal-resolution precipitation dataset, Integrated Multi-satellite Retrievals for GPM (IMERG), extratropical cyclone precipitation is evaluated in two reanalyses and two climate models. Based on cyclone-centered composites, all four models overestimate precipitation in the western subsiding and dry side of the cyclones, and underestimate the precipitation in the eastern ascending and moist side. By decomposing the composites into frequency of occurrence and intensity (mean precipitation rate when precipitating), the analysis reveals a tendency for all four models to overestimate frequency and underestimate intensity, with the former issue dominating in the western half and the latter in the eastern half of the cyclones. Differences in frequency are strongly dependent on cyclone environmental moisture, while the differences in intensity are strongly impacted by the strength of ascent within the cyclone. There are some uncertainties associated with the observations: IMERG might underreport frozen precipitation and possibly exaggerate rates in vigorously ascending regions. Nevertheless, the analysis suggests that all models produce extratropical cyclone precipitation too often and too lightly. These biases have consequences when evaluating the changes in precipitation characteristics with changes in cyclone properties: the models disagree on the magnitude of the change in precipitation intensity with a change in environmental moisture and in precipitation frequency with a change in cyclone strength. This complicates accurate predictions of precipitation changes in a changing climate.

1. Introduction

Over the past several years, considerable effort has been put into finding new ways of evaluating general circulation models so that compensating errors can be avoided and the underlying processes that might cause uncertainties can be better isolated. Among the quantities that are important for a model to represent accurately, one that stands out is

precipitation, and yet, because of the sheer number of processes involved for its production and their nonlinear nature, it is also a difficult parameter to evaluate (Tapiador et al. 2019).

Precipitation in the midlatitudes is predominantly produced in extratropical cyclones (Hawcroft et al. 2012), and therefore a number of techniques to estimate precipitation in the cyclones have been proposed. Of interest here are cyclone-centered composites. Initially introduced for cloud types by Lau and Crane (1995), they have been extensively used as well to provide information on precipitation processes in cyclones (Field and Wood 2007; Wong et al. 2018), on how these processes might evolve in a changing climate

(Bengtsson et al. 2009; Catto et al. 2011), and on how well they are represented in general circulation models (GCMs) (Bauer and Del Genio 2006; Field et al. 2008, 2011; Catto et al. 2010; Hawcroft et al. 2016, 2017; Yettella and Kay 2017; Booth et al. 2018a). The cyclone-centered view allows us to separate areas of ascent from subsidence as well as regions of cold/dry and warm/moist air. This provides a natural delineation between different precipitation conducive regimes. There have been a number of studies that have used precipitation in cyclones to evaluate model performance (e.g., Booth et al. 2018a), but none has decomposed the mean precipitation into the frequency of occurrence and mean precipitation rate when precipitating specifically in a cyclone-centered frame of reference. This distinction is important because the impact of a light and frequent rain event is different from the impact of an event with intermittent heavy rain, while producing on average the same amount at the surface. Related to this, models designed for climate change predictions need to accurately represent the sensitivity of precipitation to changes in environmental conditions, in terms of total amounts but also separately in terms of frequencies and rates.

Because compositing on cyclones includes a large number of cases at high temporal resolution, relatively short periods of observations can be used. This means that recently launched missions such as the Global Precipitation Measurement mission (GPM; Hou et al. 2014), which includes a dual-frequency precipitation radar on board the *Core Observatory* satellite (Skofronick-Jackson et al. 2017), can already be utilized. Therefore, in this study, we use the high-spatial-resolution ($0.1° \times 0.1°$) and high-temporal-resolution (30 min) Integrated Multi-satellite Retrievals for GPM product (IMERG; Huffman et al. 2017) to explore the performance of four distinct models for their representation of precipitation characteristics in extratropical cyclones. While a reanalysis is highly constrained for its thermodynamics fields through data assimilation, it relies on modeling for its representation of precipitation. Therefore, we refer to the two reanalyses examined herein as models. These reanalyses are the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2; Gelaro et al. 2017), and the European Centre for Medium-Range Weather Forecasts interim reanalysis (ERA-Interim; Dee et al. 2011). The other two models are free-running GCMs, the latest National Center for Atmospheric Research model (NCAR) Community Atmosphere Model, version 6 (CAM6; e.g., Gettelman et al. 2018), and a development version of the latest Geophysical Fluid Dynamics Laboratory (GFDL) Atmosphere Model, version

4 (AM4; Zhao et al. 2018). For the comparison, we focus on the mean precipitation in the cyclones (an average of events including those with zero precipitation rates), as well as on the frequency of occurrence of precipitation (fraction of the precipitating events out of all events) and the mean rate when precipitating (an average of only precipitating events). By decomposing precipitation this way, we can identify whether the performance of the models vis-à-vis the observations for total precipitation is affected by compensating errors in frequency and intensity. In addition to comparing models and observations, we discuss the potential uncertainties associated with the observations that might explain some of the differences. Then we investigate how changes in cyclone properties relate to changes in the precipitation characteristics for the observations and the models. We discuss the implications of our results for the model representation of the sensitivity of precipitation to changes in environmental conditions in extratropical cyclones.

2. Data and methodology

The study is focused on the latitude range 30°–60° in both hemispheres, and includes observations in all seasons from March 2014 to December 2017. Here we describe the observations, models, and the method employed to compare models to observations.

a. Observations of precipitation with the IMERG product

Precipitation observations are obtained with the version 5 IMERG "final run" product (available with a 4-month delay for research applications; Huffman et al. 2017), which provides gridded precipitation rates at a resolution of $0.1° \times 0.1° \times 30$ min up to ~65°N/S. Precipitation estimates are obtained from the merging of passive microwave radiometer observations from the Global Precipitation Measurement (Skofronick-Jackson et al. 2017) constellation of satellites. The GPM *Core Observatory* is used as a standard reference to intercalibrate the individual radiometers in the constellation. A summary of the various steps involved in the production of the IMERG precipitation estimates is given in Tan et al. (2017) along with a discussion of uncertainties at different temporal and spatial resolution, and a full description of the algorithm is available in Huffman et al. (2017). A simple wet-bulb-temperature test is applied to separate liquid from ice phase to provide a probability of liquid phase precipitation as well (Huffman et al. 2017). In Naud et al. (2018), we found a good agreement between IMERG and *CloudSat* (Stephens et al. 2002; Haynes et al. 2009) for precipitation in extratropical cyclones, with a maximum

TABLE 1. Model characteristics and estimated minimum precipitation rate for IMERG based on Tan et al. (2017) for a 1.25° × 1° final spatial resolution.

Model	Original spatial resolution	Time averaging	Epoch	Minimum precipitation rate (mm h^{-1})
MERRA-2	0.625° × 0.5°	1-h mean	2014–16	0.013
ERA-Interim	1° × 1°	6-h mean	2014–15	0.005
GFDL	1.25° × 1°	6-h mean	2008–12	0.005
CAM6	1.25° × 0.943°	Instantaneous	2010–14	0.018

difference within $0.02 \, \mathrm{mm \, h^{-1}}$. While IMERG and *CloudSat* reported similar frequencies of light, medium, and intense precipitation, the difference in precipitation rates within the cyclones was found to change in sign depending on the amount of environmental (i.e., cyclone-local) precipitable water (PW). The IMERG product reported greater precipitation rates than *CloudSat* ($\sim0.04 \, \mathrm{mm \, h^{-1}}$ difference in precipitation averaged in a 1500-km region centered on the cyclones) in moist environments (PW > 19 mm) and lower precipitation rates (up to $0.07 \, \mathrm{mm \, h^{-1}}$ difference) in dry (PW < 11 mm) and moderately moist (11 < PW < 19 mm) environments (Naud et al. 2018). That analysis provides a benchmark to evaluate the magnitude and importance of biases between models and IMERG precipitation. Overall, with its high temporal resolution and global coverage, IMERG offers a much larger sample despite its relatively shorter observing period than *CloudSat*, is more accurate [see Naud et al. (2018) for details] than the current version of the combined GPM core mission precipitation product (Grecu et al. 2016), and is better suited to track cyclones than the daily Global Precipitation Climatology Project dataset (Huffman et al. 2001). Therefore, it is the ideal dataset for this model evaluation.

b. Reanalyses and free-running GCM precipitation products

For the reanalyses, we use MERRA-2 total precipitation rates and snowfall, and ERA-Interim total precipitation rates. In Naud et al. (2018), the cyclone-centered total precipitation composites from both reanalyses were found to be very close to IMERG/*CloudSat* estimates, with the mean precipitation in a radius of 1500 km from the storm centers within $0.01 \, \mathrm{mm \, h^{-1}}$ from the observations. However, the frequency of occurrence or mean intensity (i.e., rate when precipitating) was not tested and the agreement might be the result of compensating errors.

For the free-running GCMs, we use the latest version of the NCAR model, CAM6, which is slated to be part of the next CMIP exercise [see Gettelman et al. (2018) for a description of basic physics parameterizations],

and a development version of the GFDL AM4 model [see Zhao et al. (2018) for a detailed description of the operational version, AM4.0] referred to here as devAM4 to distinguish it from the operational AM4.0 version. More details on CAM6 are also available on the development site (http://www.cesm.ucar.edu/working_groups/Atmosphere/development/). Both atmosphere-only integrations of the CAM6 and devAM4 are forced with climatological sea surface temperatures over a 5-yr period (2010–14 for CAM6, 2008–12 for devAM4) and output precipitation, sea level pressure, precipitable water, and 500-hPa vertical velocities every 6 h. CAM6 provides instantaneous surface precipitation rates whereas devAM4 provides a 6-hourly mean.

As indicated in Table 1, while the observations are available for 2014–17 at the time of this study, the free-running GCMs are available for differing periods. Because they are free running, the GCM cyclones are not coincident in time or space with those in the reanalysis. By using a 5-yr period, we ensure that we have enough cyclones to obtain a climatologically adequate representation of the cyclone precipitation, and therefore limit the impact of interannual variability (e.g., Naud et al. 2018). For the two reanalyses, because we can match the observed cyclones in space and time, shorter periods can be used. Based on the availability of reanalysis output at the time of this work, 2014–16 and 2014–15 were used instead for MERRA-2 and ERA-Interim, respectively.

c. Adjusting the precipitation products for comparison

Because we want to compare precipitation frequency of occurrence and rain rates when raining from observation products with modeled estimates, it is paramount to ensure that the spatial and temporal resolution of both precipitation products is identical. For this first step, we impose the same spatial resolution to all models and IMERG. We chose a 1.25° × 1° spatial resolution for the comparison. However, each model has its own temporal resolution and frequency of output (mean over a period or instantaneous output) and thus IMERG needs to be separately adjusted for this as well.

We therefore average IMERG products over different time scales to match the modeled precipitation: 6-hourly mean, 1-hourly mean, and instantaneous (see Table 1 for details).

The second step accounts for the limited sensitivity to light precipitation of any observational dataset. While models can technically report precipitation rates over an infinite range of intensity, observational datasets rely on instruments that have a finite detectability limit. In the case of IMERG, Tan et al. (2017) provide a useful formula to derive the minimum detectable and reliable precipitation rate for any spatial and temporal resolution after regridding. IMERG and modeled precipitation rates that are less than this minimum threshold are set to $0\,mm\,h^{-1}$ to ensure that two sets have comparable sensitivity to light rain before the comparison is conducted (see Table 1 for the threshold used for each model).

In this study, we are interested in oceanic cyclones and precipitation over the open ocean. For this, we only select cyclones with a center (location of the minimum in sea level pressure) over the ocean, but this does not prevent the surrounding area to include land or sea ice. Because IMERG does not provide precipitation information poleward of approximately 65°N/S, we impose a conservative cutoff at 60°N/S latitude in all models (i.e., none of the grid cells outside of 60°S–60°N are used, regardless of where the storm centers are found) so we find very little impact from sea ice. For the reanalyses, we only use the precipitation rates where and when IMERG is available, so the cutoff is included by construct. Using MERRA-2 for which we have information on both land and sea ice presence in each grid cell, we tested the impact on the precipitation of including versus discarding grid cells where land or sea ice is present. No impact from sea ice was found. Land can be present in various sectors of the cyclones but on average has a rather negligible impact, if any. Therefore, even when information on surface type is not available, our results are not affected by possible contamination by land surfaces.

d. Compositing method

The analysis herein is focused on extratropical cyclones, and we use the center of the cyclones as an anchor to average precipitation properties in both observations and models. To do this, we use a database of cyclone locations, obtained using the Modeling, Analysis and Prediction (MAP) Climatology of Midlatitude Storminess (MCMS) tracking algorithm (Bauer et al. 2016) applied to 6-hourly sea level pressures. Each 6-hourly cyclone snapshot is thereafter referred to as a "cyclone," regardless of when during the cyclone life this snapshot is obtained. Consequently, multiple occurrences of the same system at different times are included in the analysis. We then construct composites of total precipitation, frequency, and mean rate using the cyclone center (i.e., minimum in sea level pressure) as an anchor in a cyclone-centered equal area grid of 100-km spatial resolution. The same cyclone location database is used to construct cyclone-centered composites of precipitation using IMERG, MERRA-2, and ERA-Interim precipitation products. For the free-running GCMs, the same MCMS cyclone detection algorithm is applied to their respective SLP products.

The cyclone-centered square grid spans $\pm 1500\,km$ (north–south and east–west of the low) and each $100\,km \times 100\,km$ cell is populated by 1) the number of data points that fall in this grid cell based on the distance between the original grid cell and the center of the cyclone, 2) the number of these data points that do have a precipitation retrieval/product, 3) the number of these data points that do have a precipitation rate $> 0\,mm\,h^{-1}$, and 4) what the accumulated rate is. An example of the projection of the gridded precipitation products into a cyclone centered grid is shown in Naud et al. (2018).

Finally, using all cyclones in the database and the regridded precipitation products, we construct composites by superimposing the cyclone centers found in the two hemispheres between 30° and 60° latitude over the ocean. These composites provide in each $100\,km \times 100\,km$ grid cell: the mean precipitation P when $P \geq 0\,mm\,h^{-1}$, the mean precipitation rate when $P > 0\,mm\,h^{-1}$ (thereafter referred to as intensity), and the frequency of occurrence of precipitation (hereafter referred to as frequency), that is, the ratio of the total number of occurrences of $P > 0\,mm\,h^{-1}$ to the total number of data points with $P \geq 0\,mm\,h^{-1}$ across all cyclones. In some previous studies a rotation of the cyclones was applied prior to compositing to align specific features and alleviate the smoothing effect of averaging very disparate systems. This rotation was necessary to be able to relate surface properties to cyclone precipitation (e.g., Rudeva and Gulev 2011), to compare reanalysis to GCM dynamical features within cyclones (e.g., Catto et al. 2010), or to explore frontal features (Govekar et al. 2011; Naud et al. 2012). Here no rotation is deemed necessary as we are only interested in the overall spatial distribution of precipitation in all cyclones of all ages.

To obtain information on the cyclones themselves, we calculate cyclone-wide domain averages as in Field and Wood (2007). For this, no regridding is used because the analysis is focused on bulk characteristics. We collect gridded MERRA-2 products for the observations and reanalyses, and modeled fields for the GCMs of

precipitable water (PW) and vertical velocity at 500 hPa and average them in a circular area of 1500-km radius centered on the point of minimum SLP. For vertical velocities, we only average data points where it is ascending (negative in pressure coordinates), again in a 1500-km circle. By only including ascent, we obtain a more direct relation to cyclone strength than if the descending region (positive in pressure coordinates) was included in the average. This gives information on the environmental moisture amount available to the cyclone and its dynamical strength. To further help characterize the correspondence between the cyclone dynamics and precipitation, we also composited 500-hPa vertical velocity from each model in a cyclone-centered polar grid, using the same cyclones that are in the precipitation composites.

3. Precipitation in extratropical cyclones: Models versus IMERG

Using cyclone-centered composites of precipitation characteristics, we first compare the four models to the IMERG product.

a. Evaluation of modeled cyclone-centered mean precipitation against IMERG

All four models produce a cyclone-centered spatial distribution of precipitation similar to observed (Fig. 1): 1) a maximum ~250 km poleward and east of the cyclone center that extends equatorward and to the east in a comma shape and 2) a minimum to the west of the low pressure center. Because of the different temporal resolutions of the model data (and hence different averaging of IMERG for the comparisons; see section 2), there are variations in the magnitude of the maximum in precipitation, with slightly larger values when using 6-hourly mean than 1-h mean or instantaneous values. This is clearly visible in the IMERG composites (Figs. 1b,f,j,n), with the IMERG average used in comparison with ERA-Interim and GFDL (6-h mean) similar to one another but slightly different from the IMERG composites used for comparison with MERRA-2 (1-h mean) and CAM6 (instantaneous). There are slight differences between the IMERG composites using 6-hourly mean, because to match ERA-Interim only two years of cyclones are used whereas a whole 5-yr period is used to match the GFDL epoch length.

The composites of the difference between each model and IMERG show a consistent picture: the models slightly overestimate precipitation to the west of the low and underestimate precipitation to the east of the low when compared to IMERG. The sign of the bias changes with the sign of the mean cyclone vertical motions (i.e., subsidence to the west versus ascent to the east as indicated with the 500-hPa vertical velocity composites in Figs. 1c, 1g, 1k, and 1o). In other words, the models tend to underestimate precipitation in the region of ascent where it is relatively heavy, but overestimate precipitation in the subsidence region where it is intermittent and light. The relative difference composites (i.e., difference normalized by the IMERG composites; right column in Fig. 1) indicate that the relative bias in the subsidence region is larger than that in the ascent region. Because composites involve averaging, the possibility exists that the biases are related to differences in extreme, such as the presence of outliers in the cyclone database that might have a large impact on the differences but are not the norm. Therefore, we also examine the distributions of precipitation within the cyclones, for all cyclones, as reported with IMERG and all four models (Fig. 2). The distributions are overall similar in shape, confirming that the differences in the composites are not influenced by extremes: all four models tend to produce light precipitation more often than reported by IMERG, regardless of time averaging, and produce heavy precipitation less often. These differences were found to be significantly larger than the variability in IMERG intensity for each intensity bin. Next, we explore if the biases come from differences in the frequency of precipitation or from the precipitation intensity, or both.

b. Precipitation intensity and frequency

For models and IMERG, there is a maximum in frequency of precipitation around the center of the cyclones, and the frequency diminishes away from the center (Fig. 3). All four models tested here predict a frequency of precipitation that is larger than reported by IMERG everywhere in the cyclone area. Only CAM6 (Fig. 3m) retains the comma shape that is reported with IMERG (Figs. 3b,f,j,n); all other models predict a rather circular spatial distribution. This bias in the spatial pattern in the models may possibly be a result of the time averaging of the output. Recall that we aggregate the 30-min instantaneous IMERG data into a 6-h average, but it still starts as instantaneous data rather than an accumulated sum that is averaged. The difference in precipitation frequency is relatively smaller in the region of ascent than in the cold sector to the west of the low. This is demonstrated with both the absolute and relative difference composites (Figs. 3c,d,g,h,k,l,o,p). The region of ascent is an area where precipitation happens frequently across cyclones, whereas it is intermittent in the cold sector, and therefore the magnitude of the differences in frequency reflect these contrasting characteristics.
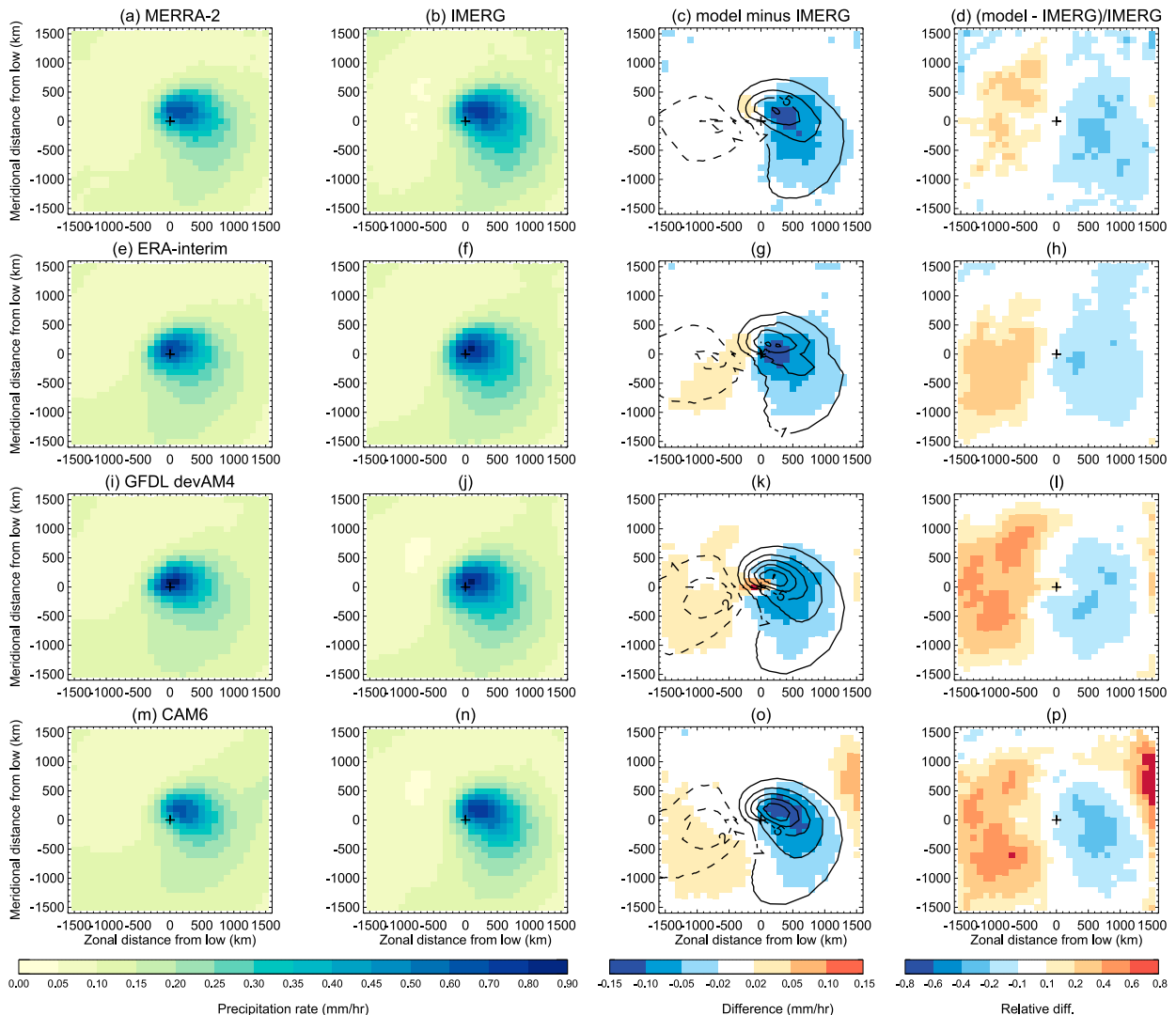
FIG. 1. (left) Cyclone-centered composites of mean precipitation rates from models, (middle left) corresponding averaged IMERG at different temporal resolution, (middle right) difference between the two, and (right) difference relative to IMERG for (a)–(d) MERRA-2, (e)–(h) ERA-Interim, (i)–(l) GFDL devAM4, and (m)–(p) CAM6. All models and IMERG are first averaged in the same global $1.25° × 1°$ resolution grid before the composites are constructed. The plus sign indicates the cyclone center; the mean includes both Northern and Southern Hemisphere cyclones, the latter having been flipped along the north–south axis to match their Northern Hemisphere counterparts. Cyclone-centered composites of 500-hPa vertical velocity are shown as black contours, with negative values set as solid contours and positive values set as dashed contours for (c) MERRA-2, (g) ERA-Interim, (k) GFDL devAM4, and (o) CAM6.

Conversely, all models predict a lower precipitation intensity than IMERG in most of the cyclone area, while retaining the comma-shape pattern in the region of maximum rates in the composite (Fig. 4). The absolute difference indicates a greater bias to the east than the west of the low (Figs. 4c,g,k,o), which is to some extent also shown by the relative difference composites, albeit with a maximum difference in intensity in the polar half of the cold sector.

Overall, while all four models are systematically giving lower intensity and higher frequencies than IMERG

in the entire cyclonic region, it appears that the difference in total precipitation in the ascent region is dominated by weaker precipitation intensity in models compared to IMERG (Fig. 4), while in the subsidence region the difference in total precipitation is dominated by a greater frequency (Fig. 3). This suggests that the differences are caused by potentially different processes; therefore, we next investigate the impact of the cyclone properties on these differences.

The cyclone properties used are 1) the cyclone-wide mean precipitable water and 2) the mean of the ascending
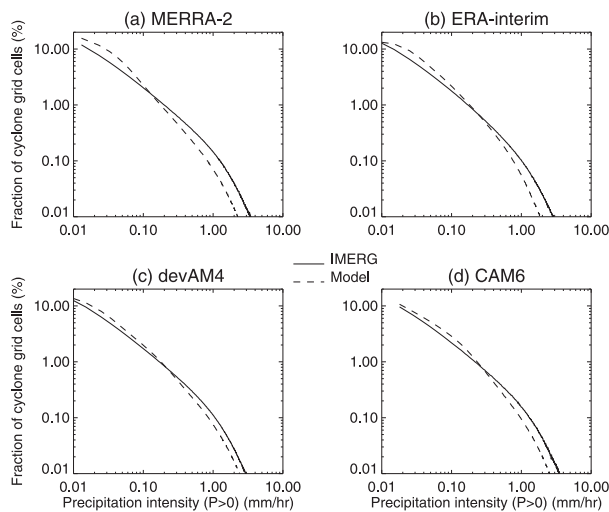
FIG. 2. Frequency distribution of precipitation intensity within cyclones, for all cyclones, according to IMERG (solid) and (a) MERRA-2, (b) ERA-Interim, (c) devAM4, and (d) CAM6 (dashed).

vertical velocities (ascent strength). Precipitation in cyclones was shown to depend on both environmental moisture amount and cyclone dynamics using surface wind speed (Field and Wood 2007; Pfahl and Sprenger 2016). Here, to characterize the cyclone strength, we choose vertical velocities in the ascent region because it relates directly to the production of precipitation as in warm conveyor belts (e.g., Eckhardt et al. 2004). Using the cyclone-centered bulk averages obtained from MERRA-2, we examined the relationship between cyclone mean ascent strength and PW, and the correlation is small ($r = -0.14$). Similarly, using all the grid points within individual cyclones we find no correlation between PW and vertical velocity at 500 hPa, for either ascending or subsiding regions. This lack of a relationship likely relates to the fact that cyclones tend to travel poleward (i.e., to drier regions) as their circulation intensifies. Booth et al. (2018b) show that this fact explains why there is a time lag between the instant of maximum in precipitation and the instant of maximum cyclone intensity for extratropical cyclone life cycles. This result is consistent with the negative correlation between PW and surface winds found by Field and Wood (2007). Given these prior results, and the weak correlation between PW and ascent strength, we use both factors to conditionally sort the cyclones.

4. The impact of cyclone properties on precipitation differences

For this part of the investigation, we focus on comparing IMERG and MERRA-2 to simplify the presentation,

but will discuss the generality of the results for the other models (figures for these models are in the supplemental material).

a. Comparison of MERRA-2 and IMERG cyclone precipitation as a function of PW and ascent strength

We classify the cyclones based on their 1500-km-radius mean PW (using 11 mm to separate dry and medium cyclones and 19 mm to separate medium and wet cyclones) and their mean ascent strength (using $-6.8\,\mathrm{hPa\,h^{-1}}$ to separate strong and moderate cyclones and $-4.7\,\mathrm{hPa\,h^{-1}}$ to separate moderate and weak cyclones). These three PW and three ascent strength categories were defined by dividing the entire cyclone database for 2006–16 into equal population subsets (Naud et al. 2017). Combining these PW and ascent strength categories we obtain nine subsets of the cyclone database and explore the impact of these two parameters on the precipitation intensity using IMERG (Fig. 5). In the ascent region, both PW and ascent strength enhance precipitation intensity as they increase, while much more subtle changes occur in the cold sector. While the impact of there being different amounts of cyclone-centered PW is mainly visible as a change in the composite-mean precipitation intensity near the center of the cyclones, changes in the ascent strength affect the size of the comma region and the rest of the warm sector as well. Similarly, frequency of precipitation increases with ascent strength in most of the cyclone area but with the maximum increase at the center (Fig. 6). It also increases at the center with increasing environmental PW, but the changes are less clear in other parts of the cyclones: there is a tendency for a region of relatively greater frequency at the tail of the comma to move from east to west as PW increases on the equator side of the low, but this is not seen for strong cyclones. Similar sensitivities are also found when using the modeled precipitation characteristics; however, there are differences that might have important implications for understanding the model biases.

First, the difference in precipitation intensity between MERRA-2 and IMERG is strongly modulated by the strength of the ascent (Fig. 7): the bias is rather uniform in most of the cyclone area, with a maximum in the region of ascent, which increases as cyclone strength increases, regardless of the amount of moisture in the cyclone environment. In fact, the sensitivity of the difference in intensity to changes in PW is not that clear. Similar results are found for the other models (see the online supplemental material). However, the *relative* difference in precipitation intensity normalized with IMERG intensity reveals no systematic changes in
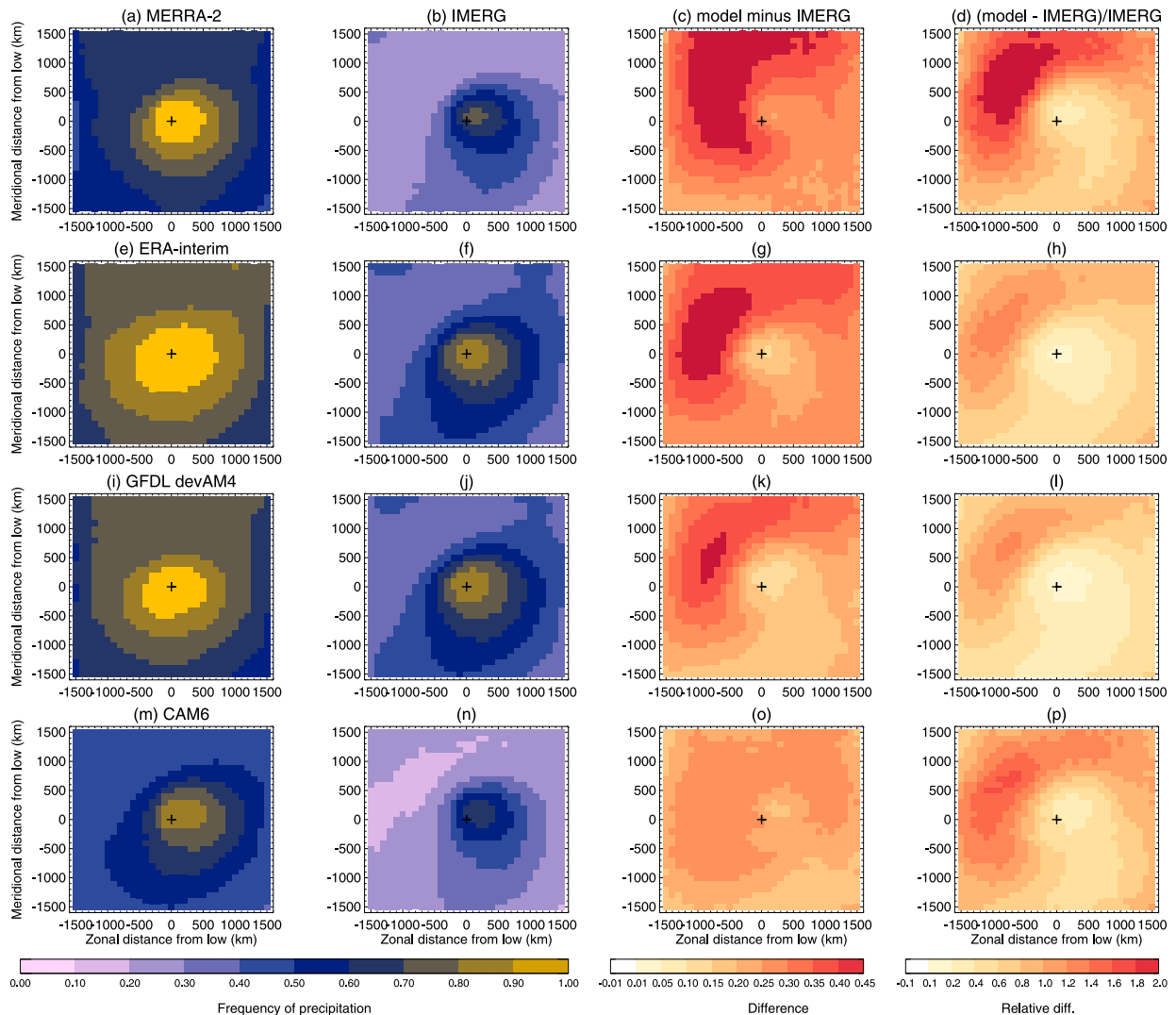
FIG. 3. (left) Cyclone-centered composite of precipitation frequency from models, (middle left) corresponding averaged IMERG, (middle right) difference between the two, and (right) difference relative to IMERG for (a)–(d) MERRA-2, (e)–(h) ERA-Interim, (i)–(l) GFDL devAM4, and (m)–(p) CAM6. The plus sign indicates the cyclone center; the mean includes both Northern and Southern Hemisphere cyclones, the latter having been flipped along the north–south axis to match their Northern Hemisphere counterparts.

bias with storm strength or PW (not shown), suggesting that, assuming IMERG intensity reporting is not biased itself, models systematically underestimate precipitation intensity, regardless of cyclone characteristics.

In contrast, the model bias in frequency of precipitation is largest in the cold sector, and the difference there decreases as PW increases (Fig. 8). Away from the region of ascent, the differences show little dependency on cyclone strength. At the center of cyclones and in the comma region, the difference depends on both PW and ascent, and decreases as both parameters increase. In fact, the difference near the center is minimized in the strongest and wettest cyclones, and maximized in weak and dry cyclones. While the bias of MERRA-2 precipitating too frequently

tends to be largest on the poleward side of the cold sector, we observe that in the case of the lowest PW category, the bias is largest at the center of the cyclones when the cyclones are weak. If IMERG detections are assumed to be correct (more on this later), these results suggest that models might overproduce precipitation in cold and dry regions of the cyclones, especially if the cyclones are weak.

These results are found to be similar for all models tested here (figures available in the supplemental material): a propensity to predict lower intensity in strong cyclones' regions of ascent, and larger frequency in cold sectors, especially in dry environments. Assuming instead that the models are in fact correct, since they all show a similar bias, these results entail that IMERG
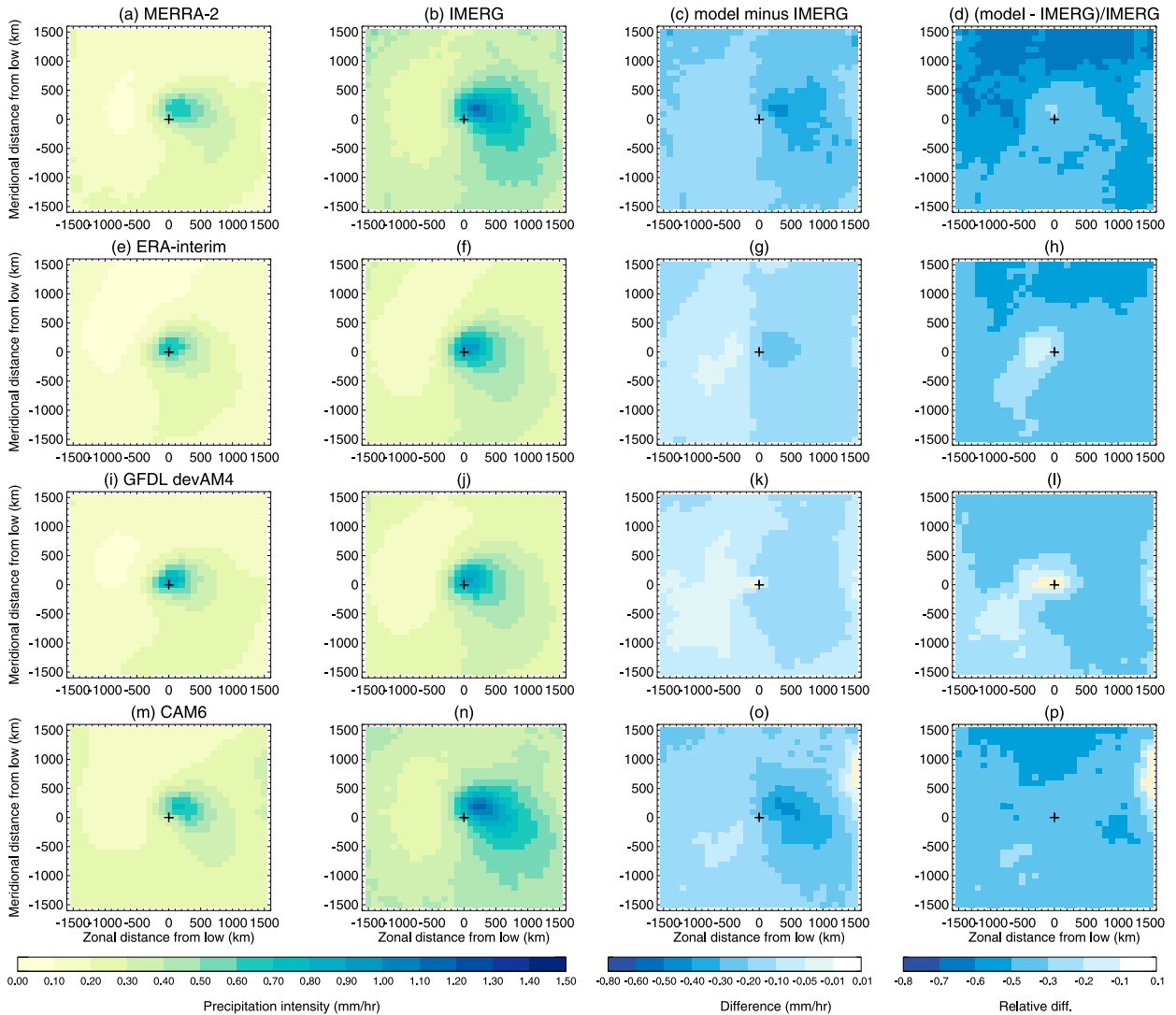
FIG. 4. (left) Cyclone-centered composites of the precipitation intensity from models, (middle left) corresponding averaged IMERG, (middle right) difference between the two, and (right) difference relative to IMERG for (a)–(d) MERRA-2, (e)–(h) ERA-Interim, (i)–(l) GFDL devAM4, and (m)–(p) CAM6. The plus sign indicates the cyclone center; the mean includes both Northern and Southern Hemisphere cyclones, the latter having been flipped along the north–south axis to match their Northern Hemisphere counterparts.

might overestimate precipitation intensity in regions where it is heavy, and underestimate frequency of precipitation in cold and dry environments. So next we discuss the possibility of such biases in IMERG.

b. Potential issues with IMERG

1) FREQUENCY OF PRECIPITATION: IMERG DETECTION SKILLS

The bulk of the measurements that are fed in the IMERG processing suite are from microwave radiometers, which have issues with light rain detection (e.g., Stephens et al. 2010). However, we ensured that the models and IMERG all use the same minimum

threshold for precipitation rate. The *CloudSat* precipitation detection product has been found to be quite effective at detecting light precipitation and is often used as a reference for other instruments and products (e.g., Behrangi et al. 2014; Stephens et al. 2010). In Naud et al. (2018) we found no evidence of light rain detection issues when we compared IMERG to *CloudSat* precipitation products in extratropical cyclones. Therefore, while we cannot eliminate the possibility that light precipitation is not always reported in the IMERG product, we also cannot eliminate the possibility that all models tend to produce light precipitation too often.

Another potential issue is with frozen precipitation. For example, precipitation rates are not retrieved
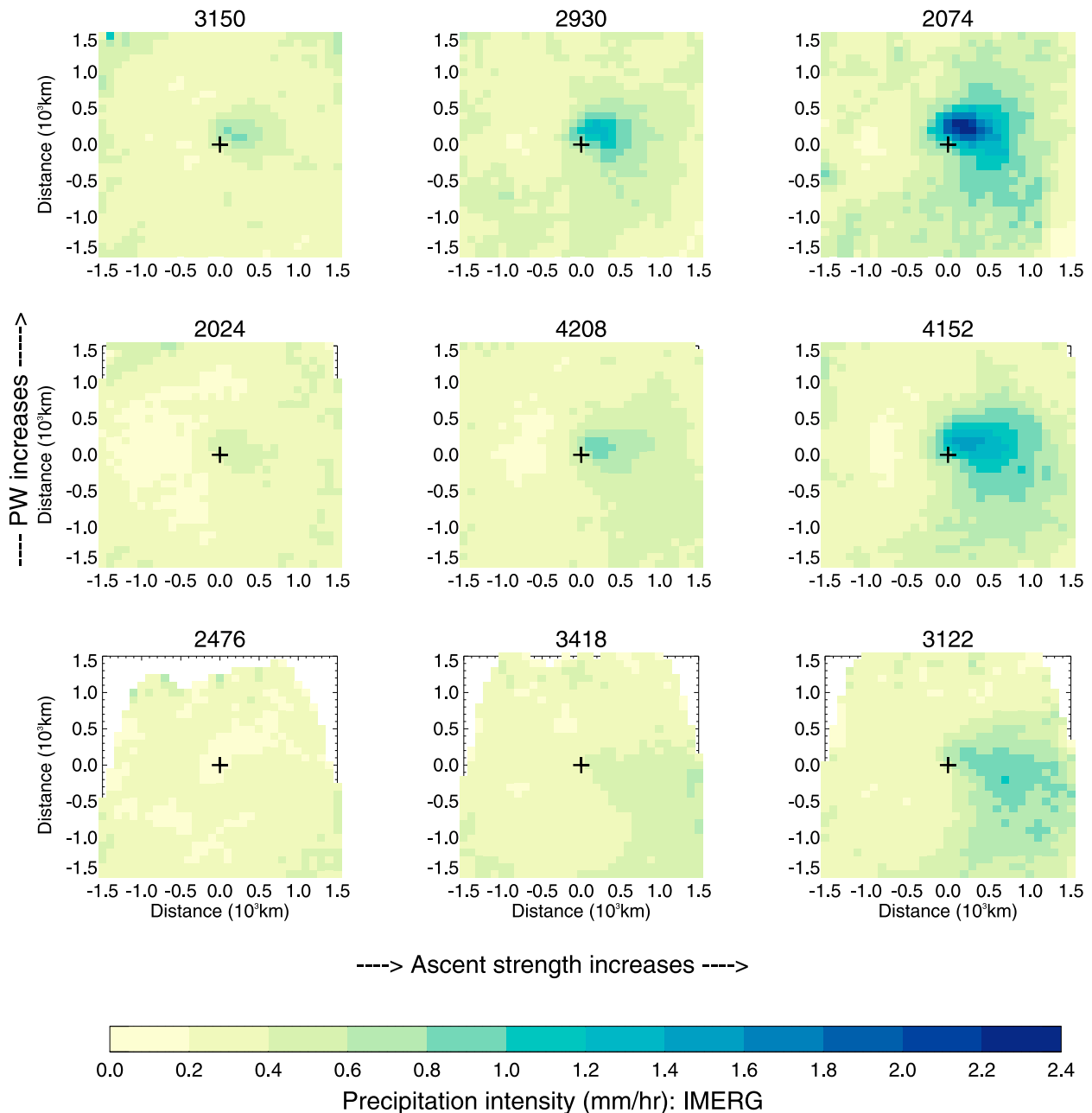
FIG. 5. Cyclone-centered composites of IMERG precipitation intensity as a function of (left to right) increasing cyclone ascent strength and (bottom to top) increasing cyclone-wide mean PW. The number above each panel is the total number of cyclones per category.

for the *CloudSat* product if the algorithm finds a bulk fraction of liquid in precipitation less than 85%. In the case of IMERG the algorithm is designed to retrieve precipitation rates regardless of the liquid fraction. But in Naud et al. (2018), we had found that in dry cyclones IMERG total precipitation was less than reported with *CloudSat*. However, to our knowledge, there has been no independent evaluation over oceans of IMERG's ability and success rate at detecting frozen precipitation.

As illustrated in Naud and Kahn (2015) for the Northern Hemisphere (see their Fig. 5), dry cyclones are preferentially situated on the poleward side of the latitude band explored here, where temperatures are lower and the melting level much closer to the surface. Therefore we performed two separate tests: 1) using MERRA-2 liquid-only precipitation (total precipitation rate minus snowfall), we test the difference in frequency of precipitation between MERRA-2 and IMERG in
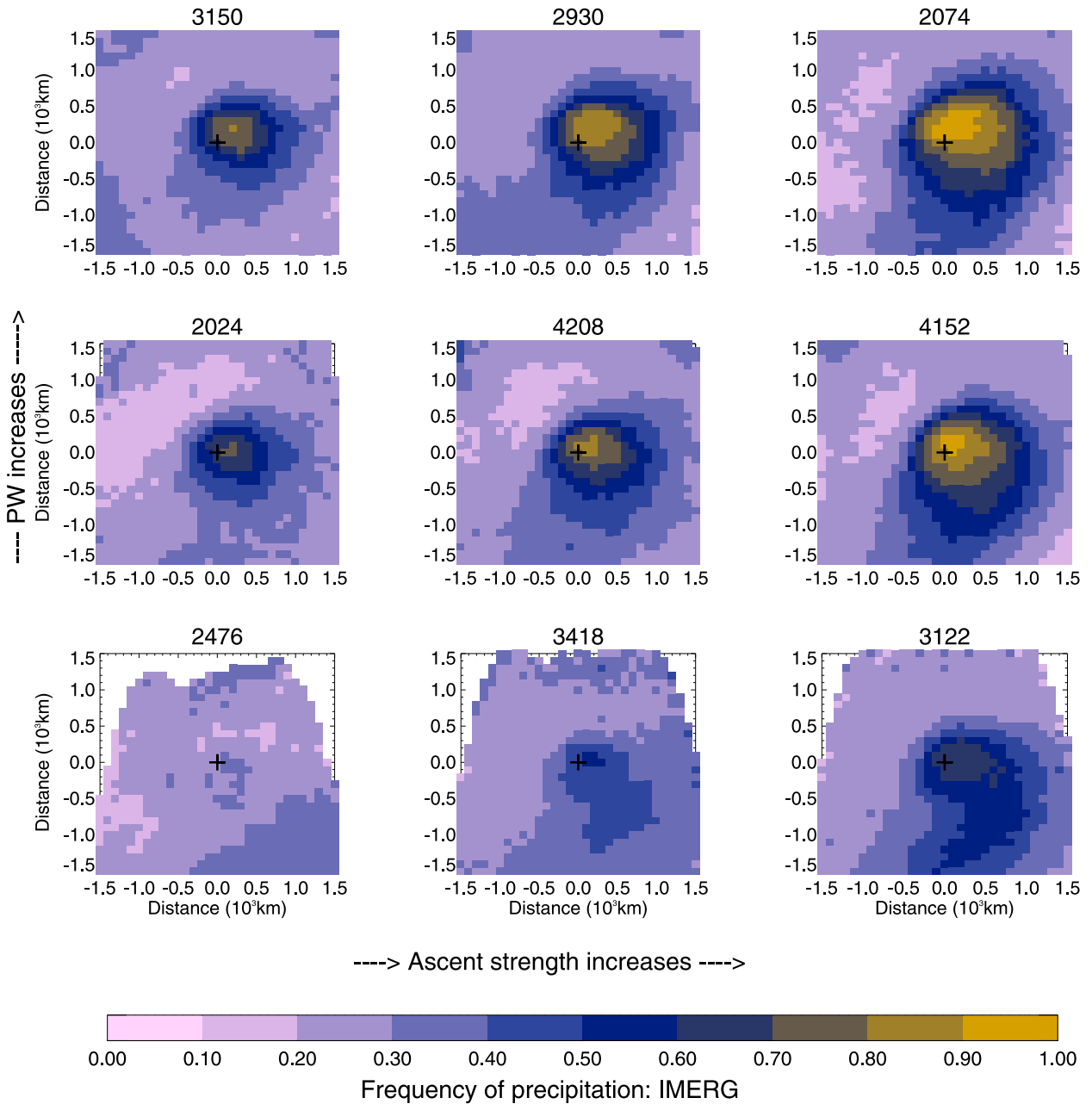
FIG. 6. Cyclone-centered composites of the IMERG frequency of precipitation as a function of (left to right) increasing cyclone ascent strength and (bottom to top) increasing cyclone-wide mean PW. The number above each panel is the total number of cyclones per category.

cyclones with PW < 11 mm (Figs. 9a–c); and 2) using IMERG's report of the probability of liquid precipitation, we build composites also for cyclones with PW < 11 mm (Figs. 9d–f). By contrasting these two series of composites and the last row of Fig. 8, we observe the following: on the polar side of the cold sector where MERRA-2–IMERG differences are the largest, removing all frozen precipitation from MERRA-2 frequencies

reduces the difference from larger than 0.5 to a negative bias of within 0.3 in a region where the probability of liquid precipitation is much less than 0.5 according to the IMERG flag. Therefore, it is quite possible that IMERG does not detect all of the frozen precipitation. However, it is also possible that MERRA-2 produces frozen precipitation too often. In fact, at least for CAM6, the correspondence between the area of maximum occurrence
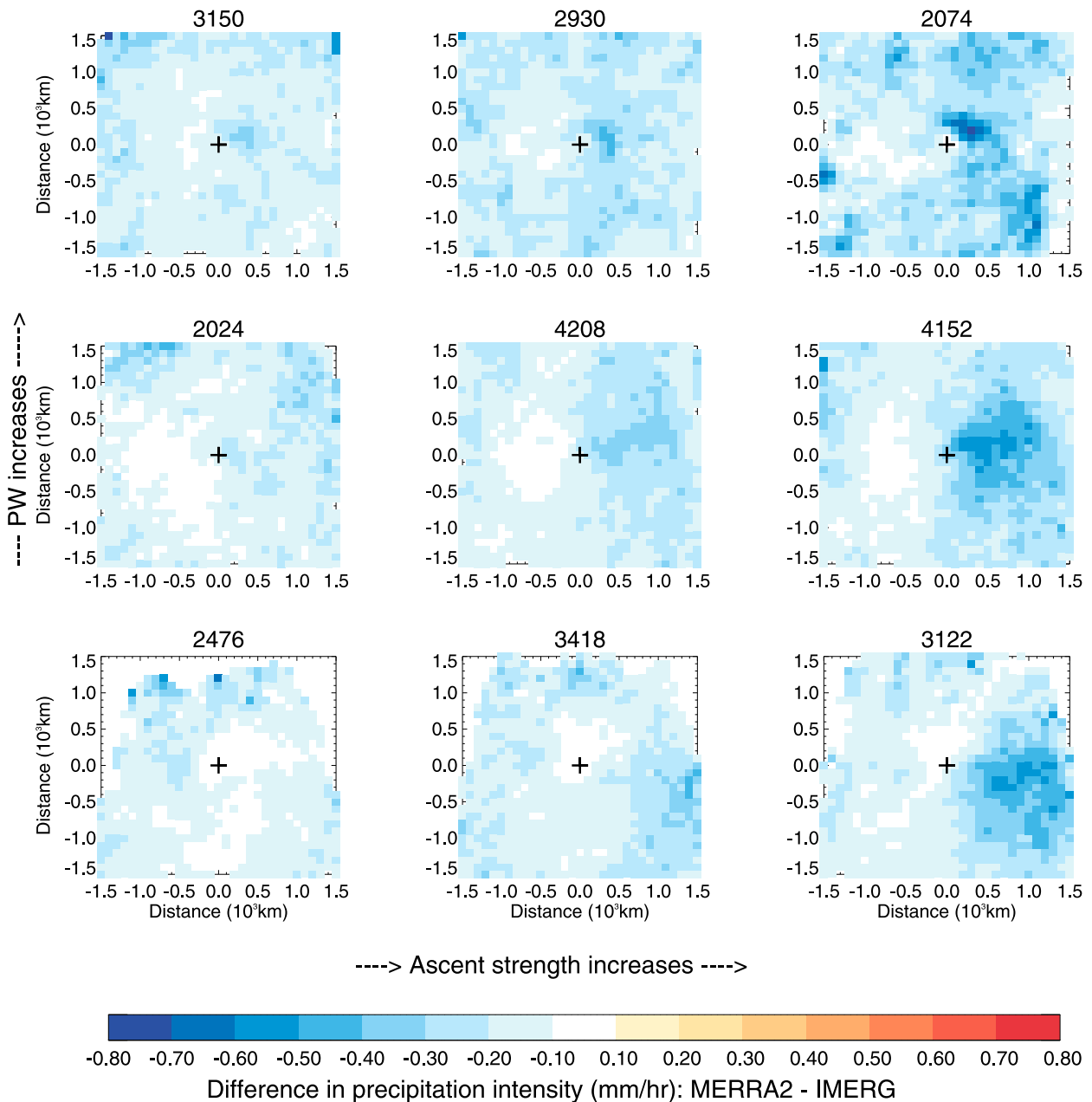
FIG. 7. Cyclone-centered composites of the difference in precipitation intensity between MERRA-2 and IMERG as a function of (left to right) increasing cyclone ascent strength and (bottom to top) increasing cyclone-wide mean PW. The number above each panel is the total number of cyclones per category for MERRA-2.

of frozen precipitation and the area of maximum difference in frequency of occurrence of precipitation is less clear (Fig. 10). Moreover, on the equatorward half of the dry cyclones, the difference in frequency of occurrence between models and IMERG cannot be explained by either issues with IMERG or models for frozen precipitation. This brings back the issue already discussed in the literature that models tend to overestimate the occurrence of light precipitation (e.g., Sun et al. 2006; Terai

et al. 2018). Our results would suggest that this issue might be exacerbated in regions of frozen precipitation.

2) THE PRECIPITATION INTENSITY DIFFERENCES: CAN IMERG OVERESTIMATE RATES?

It is quite possible that models underestimate intensity in regions of light precipitation, an issue reported elsewhere and in tune with the differences found in frequency for areas where precipitation is light and
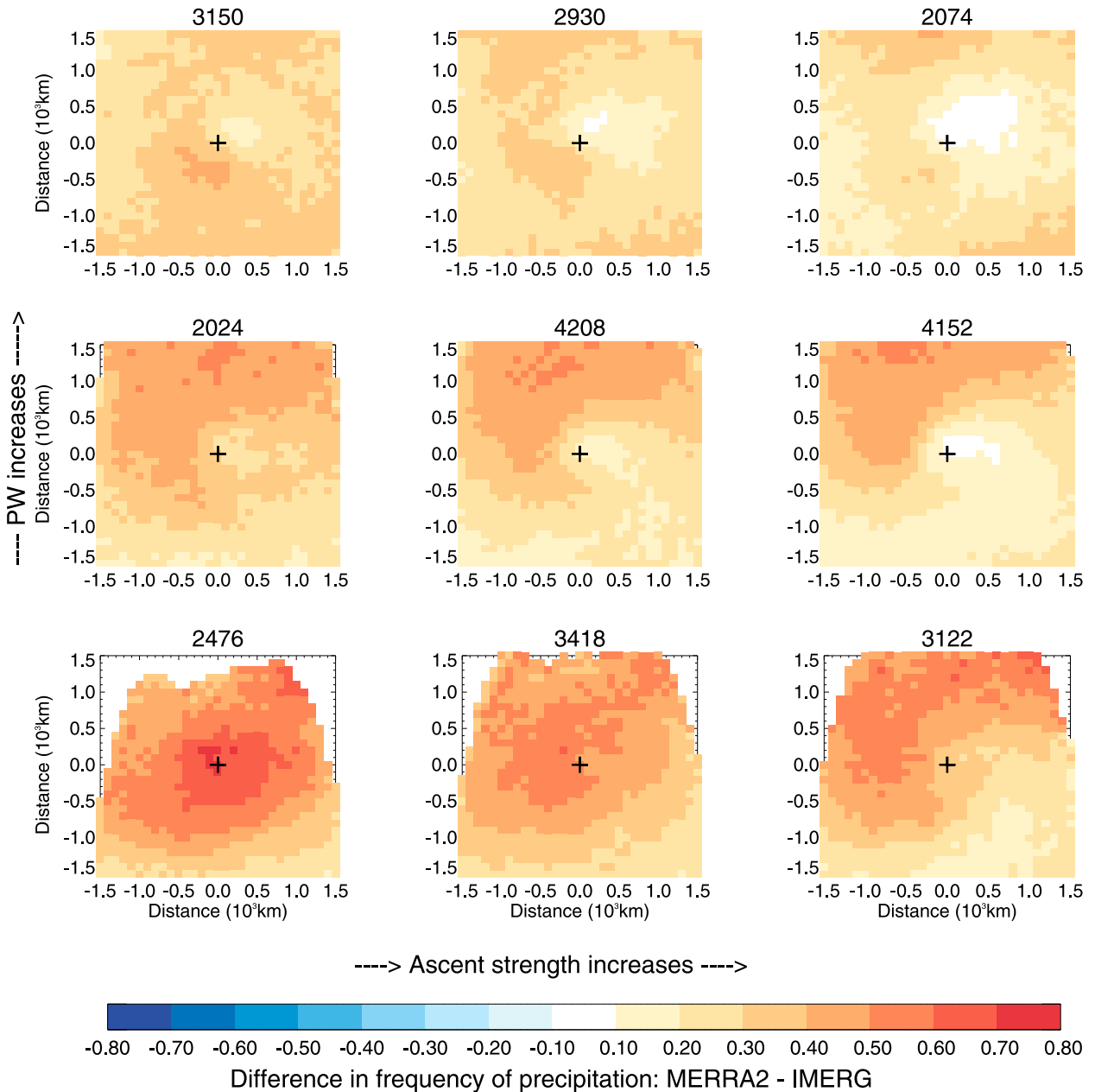
FIG. 8. Cyclone-centered composites of the difference in frequency of precipitation between MERRA-2 and IMERG as a function of (left to right) increasing cyclone ascent strength and (bottom to top) increasing cyclone-wide mean PW. The number above each panel is the total number of cyclones per category for MERRA-2.

intermittent. In contrast, in the comma region of the cyclones where precipitation rates are known to be large, we need to envisage the possibility that IMERG could overestimate precipitation intensity. It is possible if IMERG shares similar issues as the TRMM TMI products: Henderson et al. (2017) reported a tendency for precipitation rates to be overestimated in regions of heavy stratiform precipitation, which would be the dominant precipitation type in cyclone ascending regions.

However, these biases were found in the tropics and might not hold true for cooler and drier midlatitude regions. Also, while this observational uncertainty might explain the models' lower rates in the region of ascent, it does not explain their prediction of greater frequency.

c. Implications for testing sensitivities in GCMs

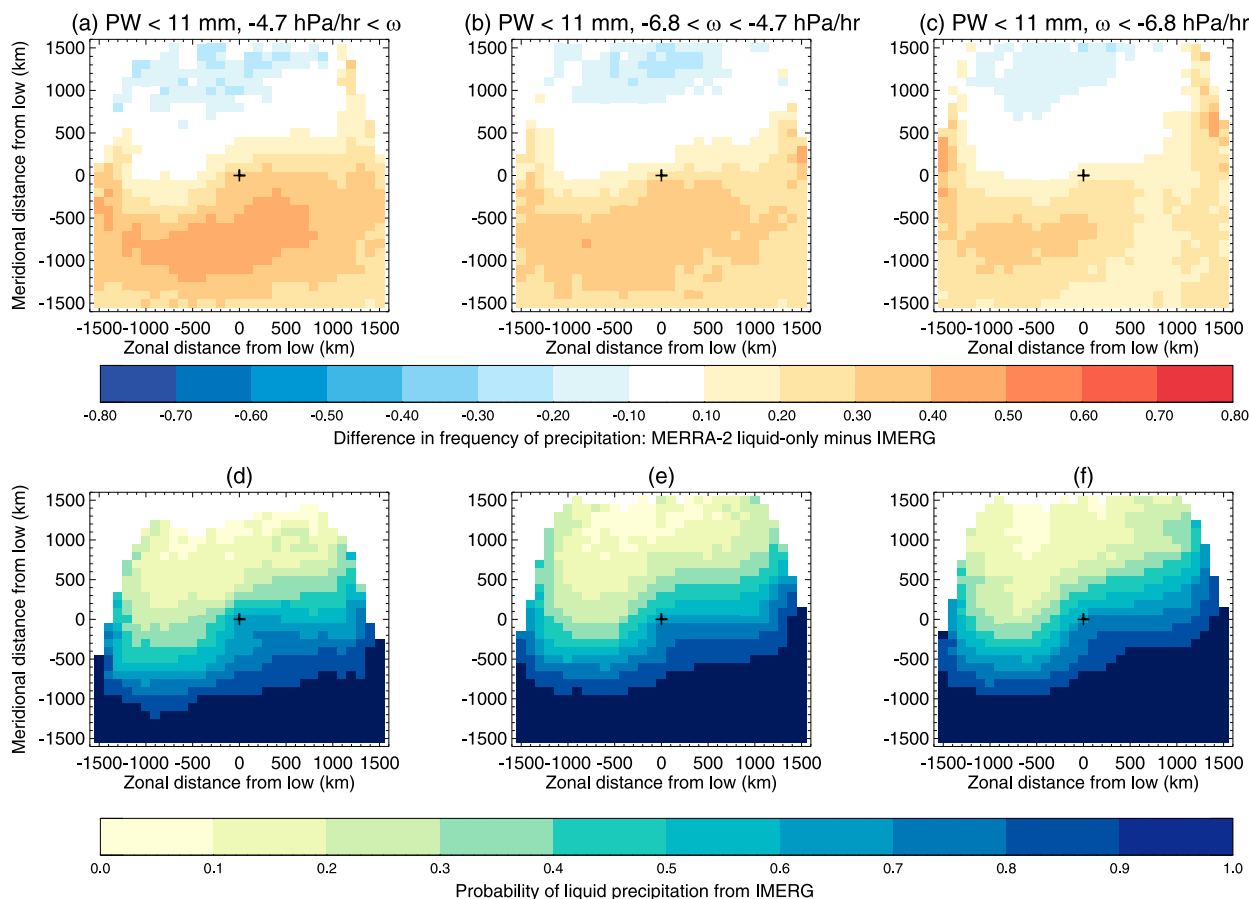As mentioned earlier, climate models are developed with the aim of reproducing a realistic sensitivity of

FIG. 9. Cyclone-centered composites of (a)–(c) the difference in frequency of precipitation between MERRA-2 liquid-only precipitation and IMERG and (d)–(f) the probability of liquid precipitation in IMERG as a function of (left to right) cyclone ascent strength for cyclone-wide mean PW < 11 mm.

various parameters to environmental changes, among them precipitation. Therefore, we would like to test whether, despite possible deficiencies in the cyclone precipitation characteristics, the models can still reproduce realistically how these characteristics change with changing environmental conditions. Here we restrict our analysis to changes in cyclone environmental PW and cyclone strength. In view of the potential shortcomings in the IMERG observations discussed in the previous section, we do not have complete confidence in the changes in precipitation intensity with changes in cyclone strength as reported with IMERG or in the changes in frequency of precipitation when cyclone PW changes. However, we can more confidently examine changes in the frequency of precipitation between strong and weak cyclones as well as the change in precipitation intensity between high and low PW cyclones. We therefore test these two changes for CAM6 and devAM4, and add MERRA-2 to test whether the differences we find are specific to climate models or also found in

reanalyses (see similar tests for ERA-Interim in the supplemental material).

Figure 11 shows the two changes mentioned above: change in frequency of precipitation as a function of a change in ascent strength and change in precipitation intensity as a function of a change in environmental PW, for IMERG at the three temporal averaging scales that match the models and the three models. Note that using a subset of cyclones to remove uncertainties caused by dry cyclones (for frequency tests) or strong cyclones (for intensity) does not change our results (see Fig S10 in the online supplemental material) so here we use the entire cyclone database.

Starting with the change in frequency for a change in ascent strength, IMERG reports a significant increase in frequency of precipitation in the comma region, but small or negative changes elsewhere, consistently regardless of time-averaging convention (Figs. 11a–c). However, the difference is largest for instantaneous reporting (Fig. 11b) than 1- or 6-h time averaging
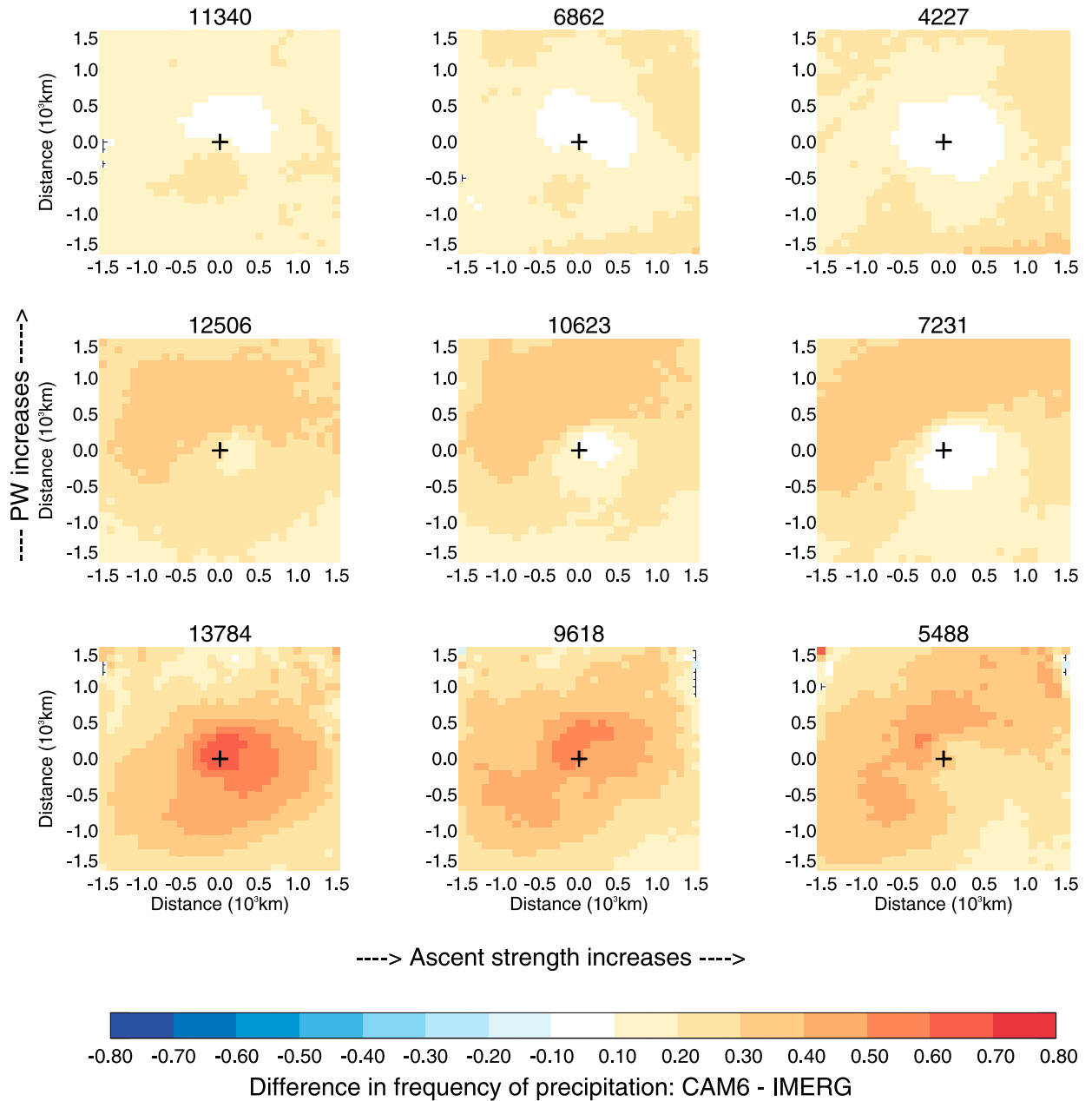
FIG. 10. Cyclone-centered composites of the difference in frequency of precipitation between CAM6 and IMERG as a function of (left to right) increasing cyclone ascent strength and (bottom to top) increasing cyclone-wide mean PW. The number above each panel is the total number of cyclones per category for CAM6.

(Figs. 11a,c). This is likely because there might be a displacement of the region of intense precipitation as cyclones travel during the hour or 6 h of the time averaging (which is performed prior to projecting in the storm centered grid and applying the minimum precipitation threshold). In this case, the frequency of precipitation can include both ascending and descending regions of the cyclone, which could potentially reduce the large number of precipitation occurrence that occurs in

ascending regions where precipitation occurs most often. For MERRA-2 the change in frequency of occurrence with a change in ascent strength is much smaller than reported with IMERG (3 times less; Fig. 11d vs Fig. 11a) and the same is found for devAM4 (Fig. 11f vs Fig. 11c). Both MERRA-2 and devAM4 also indicate a relatively larger decrease in frequency in advance and in the wake of the cold fronts (tail of the comma). In contrast, of the two GCMs, CAM6 predicts a change
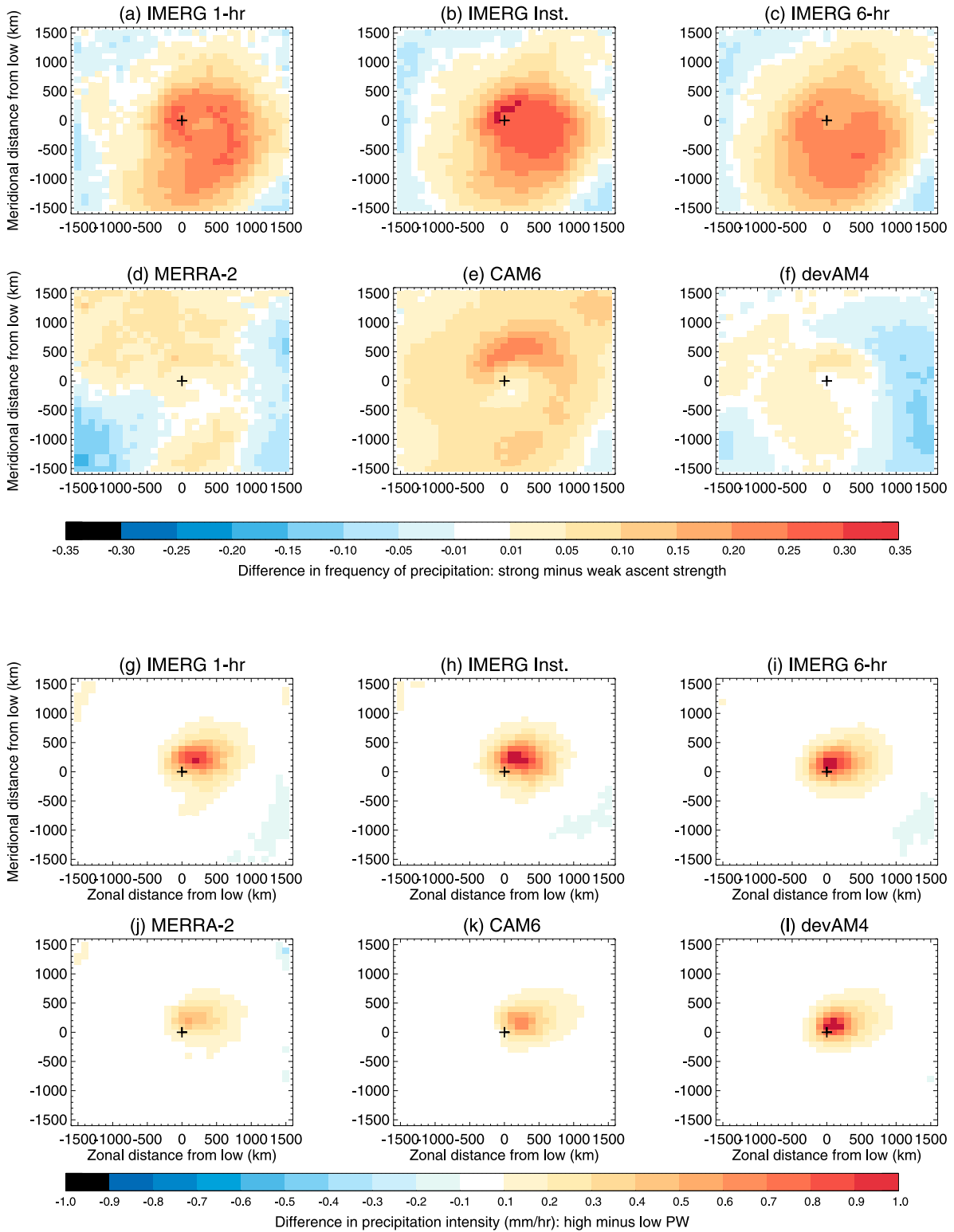
FIG. 11. Cyclone-centered composite differences in (a)–(f) frequency of precipitation between strong- and weak-ascent-strength cyclones for (a) IMERG 1-h, (b) IMERG instantaneous, (c) IMERG 6-h, (d) MERRA2, (e) CAM6, and (f) devAM4; and (g)–(l) precipitation intensity between high- and low-PW cyclones for (g) IMERG 1-h, (h) IMERG instantaneous, (i) IMERG 6-h, (j) MERRA2, (k) CAM6, and (l) devAM4.

in frequency of occurrence much closer to IMERG (Fig. 11e vs Fig. 11b), albeit still not as strong in the area at the head of the comma, while predicting mostly increasing frequencies elsewhere instead of a decrease away from the cold front. Even if IMERG is not exact, while the impact of time averaging cannot be ignored, the quite differing magnitude and spatial distribution of the response to cyclone strength among MERRA-2, devAM4, and CAM6 casts doubts on the reliability of future prediction of precipitation change in extratropical cyclones.

While the change in precipitation intensity with a change in PW is much more consistent across observations and models in terms of spatial distribution (i.e., the change affects principally the region at the head of the comma), again the magnitudes differ: MERRA-2 (Fig. 11f) and CAM6 (Fig. 11g) both predict a smaller change in intensity (about half) than that reported with IMERG (Fig. 11e), but devAM4 (Fig. 11h) predicts a much stronger change than that reported with IMERG. In contrast to changes in frequency, the change in intensity seems less dependent on changes in time averaging as instantaneous or 6-h mean IMERG are fairly similar (Fig. 11h vs Fig. 11i). Presumably this is because the mean precipitation intensity is calculated only where precipitation actually occurs, suggesting that any cyclone displacement during the hour or the 6-h period only reduces the number of precipitation occurrences, without affecting the mean intensity. Regardless of IMERG's accuracy, the intermodel spread in response of precipitation intensity to a change in cyclone PW again affects the reliability of the response of precipitation rates to future increase in PW in current models. This said, the response of precipitation in future climate simulations does not follow the same Clausius–Clapeyron relationship that is observed between PW and sea surface temperature changes (e.g., Allen and Ingram 2002; Held and Soden 2006). This entails that PW is not necessarily an ideal measure to characterize environmental changes that can affect precipitation. However, Field and Wood (2007) found that within cyclones PW and precipitation do respond similarly to changes in sea surface temperature and argue that the loss of similarity at regional or global scale comes from concurrent changes in the cyclone number or strength. So while the intermodel spread for the response to cyclone PW is problematic, the spread in response to cyclone strength might be even more so.

5. Conclusions

Using IMERG observations as a reference, we compare two reanalyses and two free-running GCMs cyclone-centered precipitation characteristics over the midlatitude oceans. While all models agree quite well with observations for total precipitation in extratropical cyclones, there are systematic differences: all models slightly underestimate precipitation in the ascent region of cyclones and overestimate precipitation in the subsidence area. When we decompose the precipitation into frequency of occurrence and precipitation rate when precipitating, we find that all four models overestimate frequency (by 10% in ascending but in excess of 100% in descending regions) and underestimate intensity (by 50% in ascending regions and up to 70% poleward of the cyclone centers) when compared to IMERG. The sign of the differences is independent of cyclone strength or environmental moisture amounts, but the magnitude changes: stronger cyclones show larger differences in precipitation intensity in the region of ascent, while drier cyclones show much larger differences in frequency of precipitation. We hypothesize that some of these differences might be caused by some limitations of the IMERG product: the product might not fully report very light and/or frozen precipitation. It is also possible that the observed intensity in strong ascent regions might be overestimated (see section 4b for an expanded discussion on potential biases in IMERG).

Nevertheless, our results are consistent with previous work by Sun et al. (2006) or Terai et al. (2018) regarding the overestimation of occurrence of light precipitation and the underestimation of heavy precipitation rates in ascending conditions. Moreover, regardless of model or observations accuracy, there are significant disagreements on the potential changes in precipitation characteristics when contrasting environmental conditions. This implies that, with the observations and models presently at our disposal, there is the potential for biases in the projections of changes in precipitation in extratropical cyclones and possibly other extreme precipitation events that affect the midlatitudes in climate change experiments. More specifically for the two reanalyses, the differences with IMERG that are found here could serve as a reminder that precipitation is modeled in these datasets and therefore caution should be exercised when using these products for climatological studies or comparisons with other models.

While model developers might find these cyclone-centered diagnostics useful to test new versions of their models during the development process, one additional test that would be invaluable is to provide similar information separately for convective and stratiform precipitation. This entails having at our disposal the means to separate convective from stratiform precipitation in the observations. Therefore, one next step will be to assess the various datasets that provide such a delineation (e.g., GPM, *CloudSat*) to provide an additional constraint for

model evaluation, as well as help explore the evolution of cyclone precipitation in a changing climate.

REFERENCES

Allen, M. R., and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrological cycle. *Nature*, **419**, 228–232, https://doi.org/10.1038/NATURE01092.

Bauer, M., and A. D. Del Genio, 2006: Composite analysis of winter cyclones in a GCM: influence on climatological humidity. *J. Climate*, **19**, 1652–1672, https://doi.org/10.1175/JCLI3690.1.

——, G. Tselioudis, and W. B. Rossow, 2016: A new climatology for investigating storm influences in and on the extratropics. *J. Appl. Meteor. Climatol.*, **55**, 1287–1303, https://doi.org/10.1175/JAMC-D-15-0245.1.

Behrangi, A., G. Stephens, R. F. Adler, G. J. Huffman, B. Lambrigtsen, and M. Lebsock, 2014: An update on the oceanic precipitation rate and its zonal distribution in light of advanced observations from space. *J. Climate*, **27**, 3957–3965, https://doi.org/10.1175/JCLI-D-13-00679.1.

Bengtsson, L., K. I. Hodges, and N. Keenlyside, 2009: Will extratropical storms intensify in a warmer climate? *J. Climate*, **22**, 2276–2301, https://doi.org/10.1175/2008JCLI2678.1.

Booth, J. F., C. M. Naud, and J. Willison, 2018a: Evaluation of extratropical cyclone precipitation in the North Atlantic basin: An analysis of ERA-Interim, WRF, and two CMIP5 models. *J. Climate*, **31**, 2345–2360, https://doi.org/10.1175/JCLI-D-17-0308.1.

——, ——, and J. Jeyaratnam, 2018b: Extratropical cyclone precipitation life cycles: A satellite-based analysis. *Geophys. Res. Lett.*, **45**, 8647–8654, https://doi.org/10.1029/2018GL078977.

Catto, J. L., L. C. Shaffrey, and K. I. Hodges, 2010: Can climate models capture the structure of extratropical cyclones? *J. Climate*, **23**, 1621–1635, https://doi.org/10.1175/2009JCLI3318.1.

——, ——, and ——, 2011: Northern Hemisphere extratropical cyclones in a warming climate in the HiGEM high-resolution climate model. *J. Climate*, **24**, 5336–5352, https://doi.org/10.1175/2011JCLI4181.1.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation systems. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Eckhardt, S., A. Stohl, H. Wernli, P. James, C. Forster, and N. Spichtinger, 2004: A 15-year climatology of warm conveyor belts. *J. Climate*, **17**, 218–237, https://doi.org/10.1175/1520-0442(2004)017<0218:AYCOWC>2.0.CO;2.

Field, P. R., and R. Wood, 2007: Precipitation and cloud structure in midlatitude cyclones. *J. Climate*, **20**, 233–254, https://doi.org/10.1175/JCLI3998.1.

——, A. Gettelman, R. Neale, R. Wood, P. J. Rasch, and H. Morrison, 2008: Midlatitude cyclone compositing to constrain climate model behavior using satellite observations. *J. Climate*, **21**, 5887–5903, https://doi.org/10.1175/2008JCLI2235.1.

——, A. Bodas-Salcedo, and M. E. Brooks, 2011: Using model analysis and satellite data to assess cloud and precipitation in midlatitude cyclones. *Quart. J. Roy. Meteor. Soc.*, **137**, 1501–1515, https://doi.org/10.1002/qj.858.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Gettelman, A., P. Callaghan, V. E. Larson, C. M. Zarzycki, J. T. Bacmeister, P. H. Lauritzen, P. A. Bogenschutz, and R. N. Neale, 2018: Regional climate simulations with the Community Earth System Model. *J. Adv. Model. Earth Syst.*, **10**, 1245–1265, https://doi.org/10.1002/2017MS001227.

Govekar, P. D., C. Jakob, M. J. Reeder, and J. Haynes, 2011: The three-dimensional distribution of clouds around Southern Hemisphere cyclones. *Geophys. Res. Lett.*, **38**, L21805, https://doi.org/10.1029/2011GL049091.

Grecu, M., W. S. Olson, S. J. Munchak, S. Ringerud, L. Liao, Z. Haddad, B. L. Kelley, and S. F. McLaughlin, 2016: The GPM combined algorithm. *J. Atmos. Oceanic Technol.*, **33**, 2225–2245, https://doi.org/10.1175/JTECH-D-16-0019.1.

Hawcroft, M. K., L. C. Shaffrey, K. I. Hodges, and H. F. Dacre, 2012: How much Northern Hemisphere precipitation is associated with extratropical cyclones? *Geophys. Res. Lett.*, **39**, L24809, https://doi.org/10.1029/2012GL053866.

——, ——, ——, and ——, 2016: Can climate models represent the precipitation associated with extratropical cyclones? *Climate Dyn.*, **47**, 679–695, https://doi.org/10.1007/s00382-015-2863-z.

——, H. Dacre, R. Forbes, K. Hodges, L. Shaffrey, and T. Stein, 2017: Using satellite and reanalysis data to evaluate the representation of latent heating in extratropical cyclones in a climate model. *Climate Dyn.*, **48**, 2255–2278, https://doi.org/10.1007/s00382-016-3204-6.

Haynes, J. M., T. S. L'Ecuyer, G. L. Stephens, S. D. Miller, C. Mitrescu, N. B. Wood, and S. Tanelli, 2009: Rainfall retrieval over the ocean with spaceborne W-band radar. *J. Geophys. Res.*, **114**, D00A22, https://doi.org/10.1029/2008JD009973.

Held, I., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *J. Climate*, **19**, 5686–5699, https://doi.org/10.1175/JCLI3990.1.

Henderson, D. S., C. D. Kummerow, D. A. Marks, and W. Berg, 2017: A regime-based evaluation of TRMM oceanic precipitation

biases. *J. Atmos. Oceanic Technol.*, **34**, 2613–2635, https://doi.org/10.1175/JTECH-D-16-0244.1.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, https://doi.org/10.1175/BAMS-D-13-00164.1.

Huffman, G. J., R. F. Adler, M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multi-satellite observations. *J. Hydrometeor.*, **2**, 36–50, https://doi.org/10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2.

——, and Coauthors, 2017: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG), Algorithm Theoretical Basis Document (ATBD) version 4.6, 28 pp., https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V4.6.pdf.

Lau, N.-C., and M. W. Crane, 1995: A satellite view of the synoptic-scale organization of cloud properties in midlatitude and tropical circulation systems. *Mon. Wea. Rev.*, **123**, 1984–2006, https://doi.org/10.1175/1520-0493(1995)123<1984:ASVOTS>2.0.CO;2.

Naud, C. M., and B. H. Kahn, 2015: Thermodynamic phase and ice cloud properties in Northern Hemisphere winter extratropical cyclones observed by *Aqua* AIRS. *J. Appl. Meteor. Climatol.*, **54**, 2283–2303, https://doi.org/10.1175/JAMC-D-15-0045.1.

——, D. J. Posselt, and S. C. van den Heever, 2012: Observational analysis of cloud and precipitation in midlatitude cyclones: Northern versus Southern Hemisphere warm fronts. *J. Climate*, **25**, 5135–5151, https://doi.org/10.1175/JCLI-D-11-00569.1.

——, ——, and ——, 2017: Observed covariations of aerosol optical depth and cloud cover in extratropical cyclones. *J. Geophys. Res.*, **122**, 10 338–10 356, https://doi.org/10.1002/2017JD027240.

——, J. F. Booth, M. Lebsock, and M. Grecu, 2018: Observational constraint for precipitation in extratropical cyclones: Sensitivity to data sources. *J. Appl. Meteor. Climatol.*, **57**, 991–1009, https://doi.org/10.1175/JAMC-D-17-0289.1.

Pfahl, S., and M. Sprenger, 2016: On the relationship between extratropical cyclone precipitation and intensity. *Geophys. Res. Lett.*, **43**, 1752–1758, https://doi.org/10.1002/2016GL068018.

Rudeva, I., and S. K. Gulev, 2011: Composite analysis of North Atlantic extratropical cyclones in NCEP–NCAR reanalysis

data. *Mon. Wea. Rev.*, **139**, 1419–1446, https://doi.org/10.1175/2010MWR3294.1.

Skofronick-Jackson, G., and Coauthors, 2017: The Global Precipitation Measurement (GPM) mission for science and society. *Bull. Amer. Meteor. Soc.*, **98**, 1679–1695, https://doi.org/10.1175/BAMS-D-15-00306.1.

Stephens, G. L., and Coauthors, 2002: The CloudSat mission and the A-Train: A new dimension to space-based observations of clouds and precipitation. *Bull. Amer. Meteor. Soc.*, **83**, 1771–1790, https://doi.org/10.1175/BAMS-83-12-1771.

——, and Coauthors, 2010: Dreary state of precipitation in global models. *J. Geophys. Res.*, **115**, D24211, https://doi.org/10.1029/2010JD014532.

Sun, Y., S. Solomon, A. Dai, and R. W. Portmann, 2006: How often does it rain? *J. Climate*, **19**, 916–934, https://doi.org/10.1175/JCLI3672.1.

Tan, J., W. A. Petersen, P.-E. Kirstetter, and Y. Tian, 2017: Performance of IMERG as a function of spatiotemporal scale. *J. Hydrometeor.*, **18**, 307–319, https://doi.org/10.1175/JHM-D-16-0174.1.

Tapiador, F. J., R. Roca, A. Del Genio, B. Dewitte, W. Petersen, and F. Zhang, 2019: Is precipitation a good metric for model performance? *Bull. Amer. Meteor. Soc.*, **100**, 223–233, https://doi.org/10.1175/BAMS-D-17-0218.1.

Terai, C. R., P. M. Caldwell, S. A. Klein, Q. Tang, and M. L. Branstetter, 2018: The atmospheric hydrologic cycle in the ACME v0.3 model. *Climate Dyn.*, **50**, 3251–3279, https://doi.org/10.1007/s00382-017-3803-x.

Wong, S., C. M. Naud, B. H. Kahn, L. Wu, and E. J. Fetzer, 2018: Coupling of precipitation and cloud structures in oceanic extratropical cyclones to large-scale moisture flux convergence. *J. Climate*, **31**, 9565–9584, https://doi.org/10.1175/JCLI-D-18-0115.1.

Yettella, V., and J. E. Kay, 2017: How will precipitation change in extratropical cyclones as the planet warms? Insights from a large initial condition climate model ensemble. *Climate Dyn.*, **49**, 1765–1781, https://doi.org/10.1007/s00382-016-3410-2.

Zhao, M., and Coauthors, 2018: The GFDL global atmosphere and land model AM4.0/LM4.0: 2. Model description, sensitivity studies and tuning strategies. *J. Adv. Model. Earth Syst.*, **10**, 735–769, https://doi.org/10.1002/2017MS001209.