DISSERTATION

SATELLITE-BASED INVESTIGATION OF CONVECTION AND PRECIPITATION IN TROPICAL CYCLONE INTENSITY CHANGE

Submitted by

Muhammad Naufal Bin Razin

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2022

Doctoral Committee:

Advisor: Michael Bell

Christian Kummerow
Kristen Rasmussen
Michael Kirby

ABSTRACT

SATELLITE-BASED INVESTIGATION OF CONVECTION AND PRECIPITATION IN TROPICAL CYCLONE INTENSITY CHANGE

Multiple hypotheses on the role of convection and precipitation in tropical cyclone intensity change have been proposed, but a scientific consensus has not yet been obtained. Some recent studies emphasize the importance of asymmetric deep convection in intensification, while others emphasize the importance of more symmetric precipitation that is not necessarily deep. To help provide clarity on this issue, this dissertation analyzes a large dataset of satellite passive microwave observations in order to obtain a sufficient sample size, which enables us to minimize the impact of the tropical cyclone environmental variability on intensity.

The first part of this dissertation involves compiling an extensive research-quality and open-access dataset of low-Earth orbit (LEO) satellite passive microwave observations centered on tropical cyclones, called the Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED). TC PRIMED consists of tropical cyclone-centric 1) inter-calibrated, multi-channel, multi-imager microwave brightness temperatures, 2) retrieved rainfall from NASA's Goddard Profiling algorithm (GPROF), 3) nearly coincident geostationary satellite infrared imagery, and 4) auxiliary data such as tropical cyclone position and intensity, ERA5 fields and derived environmental diagnostics, and satellite precipitation radar variables. TC PRIMED includes observations from over 168,000 LEO satellite overpasses of 2,101 tropical cyclones from 1998 to 2019. A simple composite analysis demonstrates the huge potential of TC PRIMED.

The second part of this dissertation uses Global Precipitation Measurement (GPM) satellite observations from TC PRIMED to train a random forest model to classify the precipitation type (i.e., convective and stratiform). The model uses raw and derived passive microwave

brightness temperature variables as input predictors and observations from the GPM dual-frequency precipitation radar as reference. This approach leverages the wider swath of the GPM passive microwave observations to obtain a larger sample size of precipitation type observation for the subsequent study on the importance of convection and precipitation in tropical cyclone intensity change. The random forest model performs very well at delineating clear air from stratiform and convective precipitation, and captures key features like the tropical cyclone eye, eyewall, and primary rainbands. However, the model struggles to detect some finer features like randomly scattered convection. Analysis of the model's performance demonstrates the importance of passive microwave texture information for precipitation type classification.

The final part of this dissertation uses the tools developed in parts one and two to investigate the distribution of convection and precipitation in tropical cyclone intensity change. This analysis minimizes the impacts of the environmental variability on intensity by only selecting observations of tropical cyclones in favorable environments. Results reveal that intensifying minor tropical cyclones have a more symmetric distribution of rainfall that is not necessarily convective, while intensifying major tropical cyclones have more numerous and intense deep convection in the upshear quadrants. The results lead to the following hypotheses: i) for minor tropical cyclones, a more symmetric distribution of rainfall is more efficient at intensifying the tropical cyclone, and ii) the occurrence of deep convection in the upshear quadrants of major tropical cyclones is optimal for intensification.

This dissertation provides the scientific community with an extensive set of tools to investigate tropical cyclones using satellite passive microwave observations and ancillary data in the form of TC PRIMED. Subsequently, the second part of this dissertation demonstrates the utility of a machine learning model in classifying multiple precipitation types from satellite passive microwave observations, with discussions on model deficiencies providing a framework to further improve such approach. Finally, this dissertation leverages a large dataset of tropical cyclone observations to clarify the role of convection and precipitation in tropical cyclone intensification.

ACKNOWLEDGMENTS

This dissertation would not have been possible without the help of my grandfather, whose love and generosity kick started my amazing adventure of studying the atmosphere here in the US. I've seen and experienced things I never would have imagined had I not embarked on this journey outside Malaysia, and for that I am eternally grateful.

I also would not have been here without the love and support of my parents. They saw my interest and passion in all things weather and, though they admittedly may not quite understood it, pushed me to pursue it. I am also incredibly grateful to have a mother who could hold me together from thousands of miles away, when all I want to do is fall apart.

I owe a great deal of gratitude to my advisor, Dr. Michael Bell, who has advised me from the very first few days of my undergraduate studies. His guidance and support were instrumental in shaping who I am as a scientist today. I can only hope that I have taken full advantage of the opportunities he has given me and continue to do so in my career.

I am indebted to so many people that I have been fortunate to have around me — too many to name here. From my sisters, extended family, friends, and the past and present members of the Bell research group. Their support have been immense, and I have learned so many things from them — including how to be a better scientist and, more importantly, a better person.

I would like to thank my committee members, Dr. Christian Kummerow, Dr. Kristen Rasmussen, and Dr. Michael Kirby for their comments and suggestions towards improving the projects presented in this dissertation.

I also owe a great deal of gratitude to the healthcare professionals in Fort Collins, who, quite literally, pieced me back together physically and mentally.

This research is supported by Colorado State University's Office of the Vice-President for Research (OVPR) under the Programs of Research and Scholarly Excellence (PRSE), the NASA Earth and Space Science Fellowship (NESSF) funding number 80NSSC18K1349, and the Office of Naval Research (ONR) funding number N000142012069.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The distribution of precipitation in tropical cyclones can impact their overall structure, leading to subsequent changes in their intensity. For example, deep convection occurring in the high inertial stability region of the tropical cyclone inner core results in the most efficient conversion of convective heating into kinetic energy, leading to intensification (Shapiro and Willoughby 1982; Schubert and Hack 1982). Whereas dominantly stratiform rainbands can produce a secondary eyewall through enhanced convergence, with this convergence occurring through persistent stratiform mid-level descending inflow (Didlake et al. 2018), or through gradient wind imbalance (Razin and Bell 2021). The secondary eyewall then initiates the eyewall replacement cycle and the ensuing intensity changes. More recently, conflicting hypotheses exist surrounding the main convective modes that are responsible for tropical cyclone intensification, with one set of studies highlighting the role of asymmetric deep convection (e.g., Guimond et al. 2010) and another set of studies highlighting the role of symmetric convection or precipitation that are not necessarily deep (e.g., Tao et al. 2017).

One of the best ways to understand the impact of precipitation on tropical cyclone intensity change is through direct observations from within the tropical cyclone itself, such as in aircraft reconnaissance missions. However, logistical limitations like the aircraft range, crew operating hours, and cost prevent frequent and global observations of tropical cyclones using aircraft. The infrequent observations mean that findings from aircraft observational studies may also depend on the environment sampled during the flight. The tropical cyclone environment is important, as environmental conditions like sea-surface temperatures, vertical shear of the horizontal winds, and tropospheric humidity strongly influence intensity. Therefore, aircraft observations are not necessarily the ideal method to obtain a statistically large sample to investigate the influence of precipitation in tropical cyclone intensity change, as the environment sampled during the infrequent reconnaissance missions may also influence the intensity change observed.

Conversely, while satellite observations are more indirect, their ubiquity results in more frequent observations of tropical cyclones around the globe. More frequent observations allow for a statistical analysis of tropical cyclones located in favorable environments, providing a more direct link between the observed precipitation structure and intensity change. Sensors aboard low-Earth orbit satellites can provide observations of the precipitation distribution in tropical cyclones. For example, the dual-frequency precipitation radar (DPR) aboard the Global Precipitation Measurement (GPM Hou et al. 2014) satellite can distinguish different precipitation types in its swath, such as convective or stratiform precipitation. However, the GPM DPR has a narrow swath (∼250 km). Therefore, even with more frequent observations of tropical cyclones, the GPM DPR can only sample a small section of the tropical cyclone. Passive microwave observations, such as from the GPM Microwave Imager (GMI) represents a compromise between the infrequent aircraft observations and the more frequent, high resolution, but narrow observations of the GPM DPR.

The GMI has a wider swath (i.e., ∼880 km) and utilizes the various radiative properties of terrestrial radiation in the microwave frequencies to deduce various atmospheric and surface properties. In the context of precipitation, observations in multiple passive microwave frequencies contain a lot of information. For example, at 37-GHz, liquid precipitation produce a higher brightness temperature compared to the radiometrically cold ocean surface. At 85–92 GHz, large precipitation-sized ice particles produce a depression in brightness temperatures. Coincident GPM DPR and MI observations allow for the development of an algorithm to infer the precipitation type using the GMI observations alone. A GMI-based precipitation type classifier will increase the sample size of observations to investigate the distribution of convection and precipitation in tropical cyclone intensity change. To use the precipitation type classifier for the GMI to investigate the distribution of precipitation in tropical cyclones, an extensive dataset of passive microwave tropical cyclone observations must first be compiled. Given the potential utility of passive microwave observations in tropical cyclone studies, passive microwave observations of tropical cyclones from other sensors will also be compiled

along with ancillary data. Such a dataset is not currently publicly available, but will hugely benefit the scientific community.

This dissertation aims to achieve three main goals. The first is to compile an extensive and public dataset of satellite passive microwave observations of tropical cyclones with ancillary data. The second goal is to develop an algorithm that uses only the GMI passive microwave observations to classify precipitation types (i.e., convective, stratiform). The ultimate goal is to use the products developed in the first two sections to analyze the distribution of precipitation during tropical cyclone intensity change. Chapter 2 discusses the dataset developed called the Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED). TC PRIMED not only contains satellite passive microwave observations, but other data as well such as retrieved precipitation, coincident infrared brightness temperatures, best-track storm data, environmental data, and re-analysis fields. Chapter 3 presents the algorithm generated to classify the precipitation type from GMI observations, while Chapter 4 presents an analysis on the distribution of precipitation in tropical cyclone intensity change using the precipitation type algorithm. Chapter 5 summarizes and concludes the work in this dissertation.

CHAPTER 2

TROPICAL CYCLONE PRECIPITATION, INFRARED, MICROWAVE, AND ENVIRONMENTAL DATASET (TC PRIMED)

2.1 INTRODUCTION

Tropical cyclones cause massive property damage and loss of life worldwide. Their significant impacts continue to motivate studies to better understand the relationship between the tropical cyclone environment, convective and precipitation structure, and intensity — all of which influence how tropical cyclone hazards such as strong winds, storm surges, and heavy rainfall affect communities. However, data availability often limits tropical cyclone studies, with many studies focused on observations or modeling of a small sample or a snapshot of tropical cyclones. In contrast to the few direct observations provided by conventional platforms, current satellites provide many indirect observations. The ubiquity of satellite data captures the breadth of the tropical cyclone life cycle in a diverse set of environments. Therefore, an open-access, long-term, and global satellite-based dataset of tropical cyclones coupled with an extensive set of ancillary data will provide researchers with the tools to conduct a more holistic study of the tropical cyclone environment, convective and precipitation structure, and intensity. Such a dataset does not exist — until now.

We have compiled a comprehensive global, satellite-based dataset of tropical cyclones with best-track, environmental diagnostics, and reanalysis data known as the Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED)[1]. TC PRIMED contains tropical cyclone-centric 1) inter-calibrated, multi-channel, multi-imager microwave brightness temperatures from low-Earth orbit satellites (LEO), 2) retrieved rainfall from NASA's Goddard Profiling Algorithm (GPROF), 3) nearly coincident infrared imagery from geostationary satellites (GEO), and 4) auxiliary data such as tropical cyclone position and intensity, ECMWF Reanalysis version 5 (ERA5) fields and derived environmental diagnostics, and satellite precipitation radar variables. With TC PRIMED's unique dataset of over 168,000

[1]Razin, M. N., C. J. Slocum, J. A. Knaff, P. J. Brown, and M. M. Bell, 2021: Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED). *Bull. Amer. Meteor. Soc.*, in peer review.

LEO satellite overpasses of 2,101 storms from 1998 to 2019, we have eliminated the tremendous effort required to ensure consistency and ease of use of the data by standardizing the different data sources and formats, and obtaining appropriate environmental and diagnostic data from a single, homogeneous source. With an all encompassing dataset, the scientific community can focus solely on the research, especially since the individual components of TC PRIMED have led to groundbreaking advancements in tropical cyclone understanding.

In the following sections, we discuss the data sources and processing steps employed to create TC PRIMED. To highlight the utility of TC PRIMED, we present a brief shear-relative composite analysis of satellite parameters. Then, we summarize this work and offer concluding remarks. For a list of key abbreviations and acronyms, we refer readers to the appendix.

2.2 Data Sources and Description

The central component of TC PRIMED is LEO observations. We prioritize LEO observations such as passive microwave brightness temperatures and precipitation radar to capitalize on their ability to directly infer various atmospheric properties such as precipitation, compared to observations in the infrared channels. However, we supplement the LEO observations with gridded geostationary infrared (IR) brightness temperatures near 11 μm, environmental diagnostic variables, and reanalysis fields. We discuss the data source for each of the TC PRIMED components below.

We obtain LEO data from the NASA Global Precipitation Measurement (GPM) constellation satellites (Hou et al. 2014). The GPM constellation consists of satellites from domestic and international partners. While each partnering agency launched their satellites with its own unique scientific and weather-monitoring objectives, the members of the partnership consolidated their data to develop a set of global microwave-based inter-calibrated brightness temperature and precipitation observations for the GPM mission. The GPM mission generates inter-calibrated microwave brightness temperatures for each constellation satellite using the GPM satellite as the "Core Observatory" for bias assessments and to provide adjustments. The inter-calibration process is complex but necessary to correct sensor biases that can be as little as 2–3 K to as large as 7–11 K (Berg et al. 2016). We use the inter-calibrated

AMSR2 & AMSR-E	AMSU-B	ATMS	GMI
10.65 V & H 18.7 V & H 23.8 V & H 36.5 V & H 89.0 V & H	89.0±0.9 QV 150.0±0.9 QV 183.31±1.0 QV 183.31±3.0 QV 183.31±7.0 QV	23.8 QV 31.4 QV 88.2 QV 165.5 QH 183.31±1.0 QH 183.31±1.8 QH 183.31±3.0 QH 183.31±4.5 QH 183.31±7.0 QH	10.65 V & H 18.7 V & H 23.8 V 36.64 V & H 89.0 V & H 166.0 V & H 183.31±3.0 V 183.31±7.0 V
MHS	**SSM/I**	**SSMIS**	**TMI**
89.0 V 157.0 V 183.31±1.0 H 183.31±3.0 H 190.31 V	19.35 V & H 22.235 V 37.0 V & H 85.5 V & H	19.35 V & H 22.235 V 37.0 V & H 91.665 V & H 150.0 H 183.31±1.0 H 183.31±3.0 H 183.31±6.6 H	10.0 V & H 19.35 V & H 21.3 V 37.0 V & H 85.5 V & H

FIG. 2.1. List of microwave frequencies in GHz and polarizations available for each sensor in TC PRIMED. V, H, QV, and QH refer to the respective vertical, horizontal, quasi-vertical, and quasi-horizontal polarizations.

microwave brightness temperatures for TC PRIMED, which ensures a long-term record of microwave brightness temperatures that are consistent with the GPM Microwave Imager (GMI) aboard the GPM Core Observatory (NASA 2021). Figure 2.1 shows a list of inter-calibrated microwave frequencies and polarizations for each sensor included in TC PRIMED.

The spaceborne radar data in TC PRIMED come from the GPM and the Tropical Rainfall Measurement Mission (TRMM; Kummerow et al. 1998) satellites. The GPM satellite carries the Dual-frequency Precipitation Radar (DPR) to measure precipitation in the Ku- and Ka-band, or 13.6- and 35.5-GHz, respectively. While the DPR has a higher spatial resolution of 5 km compared to the variable instantaneous fields-of-view (IFOV) in the various microwave channels, it has a narrower swath width than the microwave imager (NASA 2021). The GPM

DPR's predecessor, the TRMM Precipitation Radar (PR), measures precipitation only in the Ku-band and has a similarly narrower swath width compared to that of its imager.

For precipitation data in TC PRIMED, we use the NASA Goddard PROFiling algorithm (GPROF; Kummerow et al. 2015). The algorithm generates estimates of precipitation by establishing the relationship between the GMI brightness temperatures and footprint-averaged rain rates and hydrometeor profiles from the GPM DPR. The algorithm establishes this relationship by storing the set of brightness temperatures and precipitation information in a database. This database contains one year of global observations, excluding observations over snow covered surfaces. To increase accuracy and computing efficiency, the algorithm subdivides the brightness temperatures and precipitation information by surface type, total column water vapor, and 2-meter temperature from ancillary datasets. To simulate the brightness temperatures for the different constellation sensors, the algorithm uses a radiative transfer model. Therefore, each GPM constellation sensor has its own specific database of brightness temperatures and corresponding hydrometeor profiles. As a result, TC PRIMED provides precipitation estimates for each GPM constellation sensor using a common algorithm (GPROF).

The GPROF algorithm uses a Bayesian scheme to weight the database of brightness temperature and hydrometeor information. TC PRIMED users should note that Bayesian schemes reflect a statistically likely precipitation estimate, and since extreme events are naturally infrequent, GPROF generally underestimates extreme events that can occur in tropical cyclones (Brown et al. 2016). A full description of the current version of GPROF is available in the Algorithm Theoretical Basis Document (Passive Microwave Algorithm Team Facility 2018). In compiling TC PRIMED, we store the data files for the passive microwave, precipitation radar, and GPROF locally at Colorado State University and retain their original swath structure (i.e., Level-1 and Level-2; NASA Earth Data 2021).

For the gridded geostationary IR data in TC PRIMED, we use data from two sources: 1) the HURSAT dataset (Knapp and Kruk 2010) and 2) the RAMMB/CIRA IR archive (Knaff et al. 2016, 2019). The HURSAT dataset contains three-hourly, 8-km IR imagery while the RAMMB/CIRA

dataset contains 4-km IR imagery at a higher temporal resolution — both remapped to a Mercator projection. The HURSAT dataset has a much longer historical record, especially in the western North Pacific, Indian Ocean, and Southern Hemisphere basins.

For estimates of storm location and intensity, we use the tropical cyclone post-season best-track database files of the Automated Tropical Cyclone Forecast system (ATCF; Sampson and Schrader 2000). The ATCF best-track provides a smoothed record of tropical cyclone location, intensity, and other parameters from information generated by the National Hurricane Center, the Central Pacific Hurricane Center, and the Joint Typhoon Warning Center at six-hourly intervals. These operational tropical cyclone centers use the same maximum sustained 1-minute wind conventions for intensity estimation.

To represent the tropical cyclone environment, we calculate large-scale environmental diagnostic quantities in a storm-centered fashion in which we average the thermodynamic, dynamic, and kinematic fields within a predefined annulus of the storm. We base our calculations on the quantities computed for the Statistical Hurricane Intensity Prediction Scheme (SHIPS; DeMaria and Kaplan 1994; DeMaria et al. 2005). While these diagnostic quantities are readily available to the community in the form of the SHIPS Developmental Dataset, the SHIPS Developmental Dataset calculates these diagnostic quantities from the operational version of the Global Forecast System or Climate Forecast System Reanalysis (SHIPS 2020). Therefore, the developmental version contains inhomogeneities due to model upgrades. To ensure consistency throughout the TC PRIMED period, we use the fifth generation ECMWF Reanalysis product (ERA5; Hersbach et al. 2020) to calculate the environmental diagnostics at each synoptic time (00, 06, 12, 18 UTC) throughout each storm's life cycle. We obtain the reanalysis output from the European Union's Copernicus Climate Change Service on pressure- and single-level surfaces at a uniform 0.25-degree horizontal grid spacing. In addition to calculated diagnostic quantities, we also include pertinent gridded ERA5 fields in TC PRIMED.

2.3 Processing Steps

We organize TC PRIMED files by year, basin, and storm directories. In each storm directory, two types of files reside: 1) an overpass file and 2) an environmental file. An overpass

file contains the satellite-specific variables associated with one satellite overpass, whereas an environmental file contains the best-track information, environmental diagnostics, and gridded ERA5 variables at each synoptic time throughout the storm's life cycle. Therefore, for each storm, multiple overpass files exist, of which the exact number depends on various factors such as the storm duration and the number of satellites in orbit. Conversely, there would be only one environmental file for each storm. Both file types are in NetCDF format and meet the Climate and Forecast Conventions and Metadata version 1.7 (CF-1.7) and Attribute Convention for Data Discovery version 1.3 (ACDD-1.3) metadata conventions.

2.3.1 Overpass Files

We begin compiling TC PRIMED upon the release of the post-season and final best-track data by selecting a satellite overpass if its passive microwave swath traversed within 750 km and ±3 hours of the six-hourly best-track storm location. We chose the 750 km distance to focus our dataset on convective and precipitation features that are likely to be directly related to the tropical cyclone. Then, we linearly interpolate the storm's center position to the time of the closest satellite scan. After which, we calculate the satellite coverage within a 750 km radius around the interpolated storm center and retain overpasses with at least 50% coverage. We use the same method and distance criteria to select the swaths for the TRMM PR and GPM DPR by limiting the swath width of the respective TMI and GMI scans. We exclude files with irregularly-spaced fill values in the latitude and longitude arrays, which can occur due to satellite navigation uncertainties. Finally, we adopt a small change for the TRMM PR and GPM DPR rain type classification. The original rain type classification algorithm classifies shallow rain detected by the precipitation radars as convective (Awaka et al. 2016). Shallow rain is precipitation with a storm top that is lower than 1 km below the 0°C level, which Awaka et al. (2016) estimated from a reanalysis output. Since shallow and non-shallow rain can have vastly different vertical extents and passive microwave signatures, we separate shallow rain from the convective rain type in TC PRIMED to produce five main rain type categories: no rain, stratiform, convective, other, and shallow rain.

When including IR data for a particular storm, we first determine if the RAMMB/CIRA imagery archive contains any imagery of the storm. If the RAMMB/CIRA archive contains imagery of the storm, we search that archive and select IR imagery within 3 hours of the microwave overpass. If the RAMMB/CIRA archive does not contain any imagery of the storm, we search and select IR imagery within 3 hours of the microwave overpass from the HURSAT archive instead. By prioritizing the RAMMB/CIRA IR archive, we include the IR data with the highest spatial and temporal resolution for each overpass in TC PRIMED. In addition, we ensure consistency by selecting IR imagery from a single source for a particular storm.

TRMM PR and GPM DPR swaths are not always available within 750 km of the storm center even if their corresponding microwave swaths are within 750 km radius of the storm center. When these data are not within 750 km of the storm center, we do not include the PR or DPR data in the final overpass file. Similarly, we leave the IR section of the file empty if no IR data exist within 3 h of the overpass or when no IR data exist at all, which often occurs for the early parts of a storm or storms that were added to the best-track following the season. We mark these cases accordingly with a status flag in the overpass file. Table 2.1 shows the list of variables that can be found in an overpass file.

2.3.2 Environmental Files

After the processing of the overpass files, we add to the environmental file a list of basic information that includes all of the overpass file names, sensor names, satellite names, and satellite coverage fractions for each storm. We then include metadata about the storm gathered from the post-season best-track ATCF database file and calculate the environmental diagnostics from the ERA5 fields. Slocum et al. (2021) provides a more technical description of the environmental diagnostics calculations using ERA5. However, we summarize the key steps here. To calculate the storm-relative diagnostics, we re-grid the ERA5 reanalysis fields onto a 0- to 1500-km cylindrical grid with 5-km radial by 5-degree azimuthal grid spacing using bilinear interpolation. We center the grid on the tropical cyclone center, which we determined using one of two ways: 1) the tropical cyclone center in ERA5 or 2) the best-track center when the system is ill-defined (e.g., disturbance, open wave, extratropical transition).

TABLE 2.1. List of variables and their units that are available in each overpass file. Note that precipitation radar variables are available only on the TRMM or GPM satellite. The list of microwave brightness temperatures for each sensor is shown in Fig. 2.1.

Group	Variable	Units
Overpass Metadata	time	
	storm latitude	degrees north
	storm longitude	degrees east
	sensor name	
	satellite name	
	overpass number	
	coverage fraction	%
	closest synoptic time	
GPROF	surface precipitation rate	mm hr^{-1}
	convective precipitation rate	mm hr^{-1}
	vertical profile of rain water	g m^{-3}
	column-integrated ice water	g m^{-2}
	vertical profile of snow water	g m^{-3}
	surface type	
Passive Microwave	microwave brightness temperatures (See Fig. 2.1)	K
Precipitation Radar (TRMM or GPM)	status flag	
	vertical profile of reflectivity	dBZ
	surface precipitation rate	mm hr^{-1}
	precipitation type	
	surface type	
Infrared	status flag	
	time	
	brightness temperature near 11 μm	K

We locate the ERA5 tropical cyclone center using a Limited-memory Broyden–Fletcher–Goldfarb–Shanno optimization algorithm to find the minimum in the mean sea-level pressure and 850-hPa geopotential height within a 4- × 4-degree box centered on the best-track center. In cases where we cannot determine the tropical cyclone center in ERA5, such as when the best-track position is over land or when the storm's maximum intensity is less than 34 knots (17.5 m s^{-1}; 1 kt = 0.514 m s^{-1}), we use the best-track center position. Using a quality flag, we specify the method used to obtain the tropical cyclone center for the diagnostics calculation at a particular synoptic time. Finally, we include pertinent gridded ERA5 fields within 20° latitude × 20° longitude of the storm centers in the environmental file. Table 2.2 shows the list of variables that can be found in an environmental file.

TABLE 2.2. List of key variables and their units that are available in each environmental file. The list is non-exhaustive and excludes variables such as data quality flags.

Group	Variable	Units	Variable	Units
Storm Metadata	time		best-track intensity	kt
	best-track storm latitude	degrees north	best-track minimum sea-level pressure	hPa
	best-track storm longitude	degrees east	storm distance to land	km
Overpass Metadata	overpass times		overpass sensors	
	overpass file names		overpass satellites	
	overpass storm latitude	degrees north	overpass number	
	overpass storm longitude	degrees east	overpass coverage fraction	%
Diagnostics (SHIPS)	minimum central sea-level pressure	hPa	vertical profile of temperature	K
	sea surface temperature	K	warm-core temperature anomaly	K
	2-m temperature	K	vertical profile of relative humidity	%
	2-m dewpoint temperature	K	vertical profile of specific humidity	$kg\ kg^{-1}$
	10-m u-wind component	$m\ s^{-1}$	vertical profile of u-wind component	$m\ s^{-1}$
	10-m v-wind component	$m\ s^{-1}$	vertical profile of v-wind component	$m\ s^{-1}$
	shear magnitude	$m\ s^{-1}$	vertical profile of tangential wind component	$m\ s^{-1}$
	shear direction	$^{\circ}$	vertical profile of vorticity	s^{-1}
	precipitable water	$kg\ m^{-2}$	vertical profile of divergence	s^{-1}
Gridded ERA5	precipitable water	$kg\ m^{-2}$	3-D temperature	K
	large-scale rain rate	$kg\ m^{-2}\ s^{-1}$	3-D relative humidity	%
	convective rain rate	$kg\ m^{-2}\ s^{-1}$	3-D specific humidity	$kg\ kg^{-1}$
	sea-surface temperature	K	3-D u-wind component	$m\ s^{-1}$
	mean sea-level pressure	hPa	3-D v-wind component	$m\ s^{-1}$
	2-m temperature	K	3-D vorticity	s^{-1}
	2-m dewpoint temperature	K	3-D divergence	s^{-1}
	10-m u-wind component	$m\ s^{-1}$		
	10-m v-wind component	$m\ s^{-1}$		

2.4 Scientific Application

TC PRIMED consists of over 168,000 LEO overpasses of 2,101 storms that occurred between 1998 to 2019. By combining the multi-sensor and inter-calibrated microwave imagery, retrieved precipitation, IR imagery, and consistently-calculated tropical cyclone environmental diagnostics into a single publicly available dataset, TC PRIMED opens up the avenue for the scientific community to conduct new research without being limited by satellite, sensor, channel, or data consistency and availability.

To highlight the value of TC PRIMED, we present an application in the context of Typhoon Sepat (2007). Figure 2.2 shows observations from a TRMM overpass of Typhoon Sepat at 02:23:17 UTC on 17 August 2007, as it was undergoing an eyewall replacement cycle off the coasts of the Philippines and Taiwan. Note that while the satellite variables are available within a 750- × 750-km box around the interpolated storm center, Fig. 2.2 depicts a closer look of the storm to better show the tropical cyclone structures.

From this TRMM satellite overpass, researchers can obtain precipitation information from either GPROF or the TRMM PR, with the surface precipitation rate from both products shown in Figs. 2.2a and 2.2b, respectively. Note that the GPROF surface precipitation rate is lower than the TRMM PR surface precipitation rate because GPROF creates an estimate using a Bayesian averaging scheme over a larger IFOV. However, the GPROF swath is larger than the TRMM PR swath and can observe the rest of the tropical cyclone beyond the TRMM PR swath. The TRMM PR reflectivity field at 2 km altitude shows Typhoon Sepat's decaying inner eyewall, with its secondary eyewall marked by a high ring of reflectivity (Fig. 2.2c).

To investigate the distribution of convection in Typhoon Sepat, researchers can use the various satellite proxies available in TC PRIMED. For example, in the 37-GHz microwave channel, warm rain processes produce warmer brightness temperatures (TBs) relative to a radiometrically colder ocean surface. Therefore, Fig. 2.2d shows the location of warm rain processes in Typhoon Sepat. However, researchers must use caution as land can also produce warmer TBs in this channel due to its higher emissivity relative to the ocean.

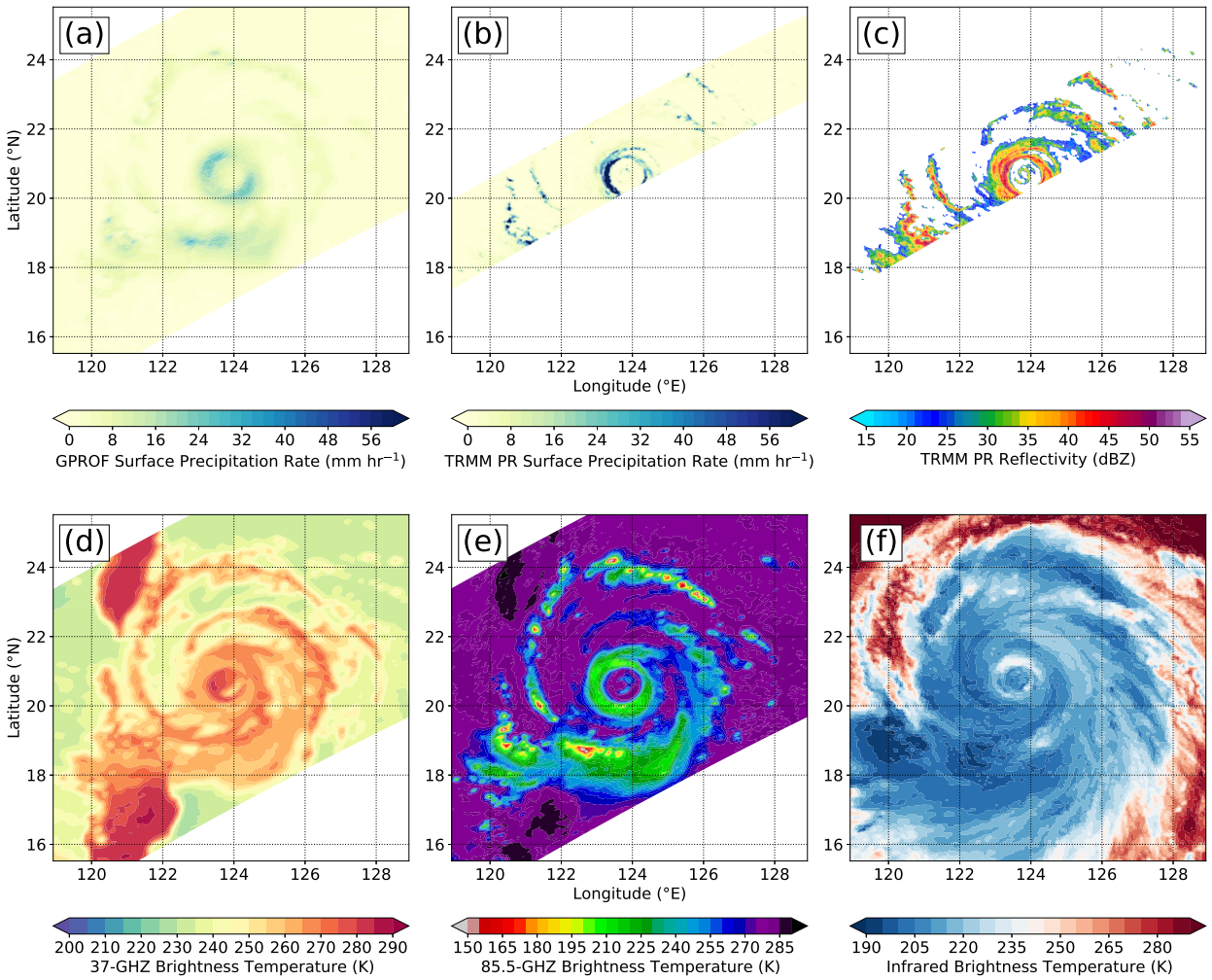

FIG. 2.2. Example of select satellite variables available in a single satellite overpass file in TC PRIMED, taken from a TRMM overpass of Typhoon Sepat (2007) at 02:23:17 UTC on August 17th 2007. (a) GPROF surface precipitation rate, (b) TRMM PR surface precipitation rate, (c) TRMM PR radar reflectivity at 2 km altitude, (d) 37-GHz microwave brightness temperatures, (e) 85.5-GHz microwave brightness temperatures, and (f) infrared brightness temperatures near 11 μm.

In the 85.5- to 91-GHz passive microwave channels, low TBs often indicate scattering of terrestrial radiation by large precipitation-sized ice particles produced in deep convection (e.g. graupel). In the IR, low TBs indicate the presence of cold cirrus ice crystals at cloud top, which are often produced by deep convection. However, cirrus ice crystals are small and do not usually produce precipitation that reaches the surface. Therefore, the low TBs in the microwave ice-scattering channels collocate more closely with deep convection, as the swirling winds of the tropical cyclone advect cold cirrus ice crystals away from the source convection.

The application of both ice-scattering and IR observations in Typhoon Sepat is shown in Figs. 2.2e and 2.2f.

While Fig. 2.2 demonstrates TC PRIMED's capability to provide a multi-variable view of the storm from just a single file, we demonstrate how the dataset can also be used to generate more extensive composite analyses to provide a long-term, globally consistent picture of the distribution of precipitation and convection in tropical cyclones. The composite analysis includes mean GPROF surface precipitation rate, polarization-corrected TBs (PCT) in the ice-scattering channels (85.5- to 89.0-GHz, hereafter 89-GHz for brevity), and IR TBs. PCTs combine the horizontal and vertical polarization TBs to remove ambiguities produced by varying surface emissivities. We use the equation derived and coefficients recommended in Cecil and Chronis (2018) for the 89-GHz PCT.

We composite storms that 1) reached at least category 1 intensity on the Saffir–Simpson hurricane wind scale (65 kt; 33 m s^{-1}) in their lifetime, 2) were at least 200 km away from land and major islands at the time of the passive microwave overpass, and 3) occurred in the North Atlantic, eastern North Pacific, central North Pacific, or western North Pacific ocean basins. For each passive microwave overpass, we bin the passive microwave and IR observations in 1) six-hourly intervals within ±72 hours of the first occurrence of a storm's lifetime maximum intensity and 2) into shear-relative quadrants: upshear right (UR), downshear right (DR), downshear left (DL), and upshear left (UL) following Corbosiero and Molinari (2002). We linearly interpolate to the passive microwave observation time the distance to land values from the best-track section of the file and the shear heading values from the diagnostics section. However, we make no distinction between the different shear magnitudes at the different observation times, nor do we make a distinction between the different intensity trends prior to and after the first occurrence of lifetime maximum intensity.

For the radial bins, we adopt a radius scaling to account for the differences in tropical cyclone size, which we calculate using a scaling parameter from Knaff et al. (2014, 2017). The scaled radius derives from azimuthally-averaged IR TB patterns, storm intensity, latitude, and climatological storm size parameters. The scaled radius of a storm is larger than its physical

radius for a storm with a climatologically small wind field. Conversely, the scaled radius of a storm is smaller than its physical radius for a storm with a climatologically large wind field. We linearly interpolate the scaling parameter from its three-hourly data points to the passive microwave observation time and bin the satellite variables in 50 km scaled radius bins (R*). We composite only storms with maximum intensity above 34 knots at the observation time, since weaker storms tend to have an extremely large scaled radius relative to its physical radius.

To ensure consistency in the GPROF surface precipitation rate and the 89-GHz TBs, we use sensors with similar 89-GHz IFOVs, such as the AMSR-E, AMSR-2, TMI, and GMI sensor. We initially composite the 8-km HURSAT and the 4-km RAMMB/CIRA IR TBs separately. However, both composites produce the same conclusion. Therefore, we combine the IR TBs from the two sources to produce one composite analysis of IR TBs. Figures 2.3, 2.4, and 2.5 show the respective shear-relative mean and standard deviation of surface precipitation rate, 89-GHz TBs, and IR TBs relative to the tropical cyclone lifetime maximum intensity.

The mean rate and radial extent of precipitation gradually increase prior to lifetime maximum intensity in all quadrants, except UL where the radial extent is constant (Fig. 2.3). At lifetime maximum intensity, the mean precipitation rate maximizes in each quadrant. The largest radial extent of precipitation occurs in DL, followed by DR, UL, and UR. After lifetime maximum intensity, mean rate and radial extent of precipitation continue to be large in the DL quadrant, with diminishing mean rate and radial extent in the DR, UL, and UR quadrants.

The evolution of convection relative to lifetime maximum intensity is similar to the evolution of precipitation, as denoted by the colder TBs in the 89-GHz and IR channels. The mean TBs grow increasingly colder while the radial extent of colder TB values expands in all quadrants prior to lifetime maximum intensity (Figs. 2.4 and 2.5). At lifetime maximum intensity, the intensity and radial extent of convection reaches a maximum. The location and intensity of convection are better shown in the 89-GHz TBs compared to the IR TBs due to the advection of cirrus ice crystals away from the primary location of convection by the tropical cyclone swirling winds. As previously mentioned, the TBs in the 89-GHz channel are insensitive to the cirrus ice crystals and are instead associated with precipitation-sized ice particles in strong

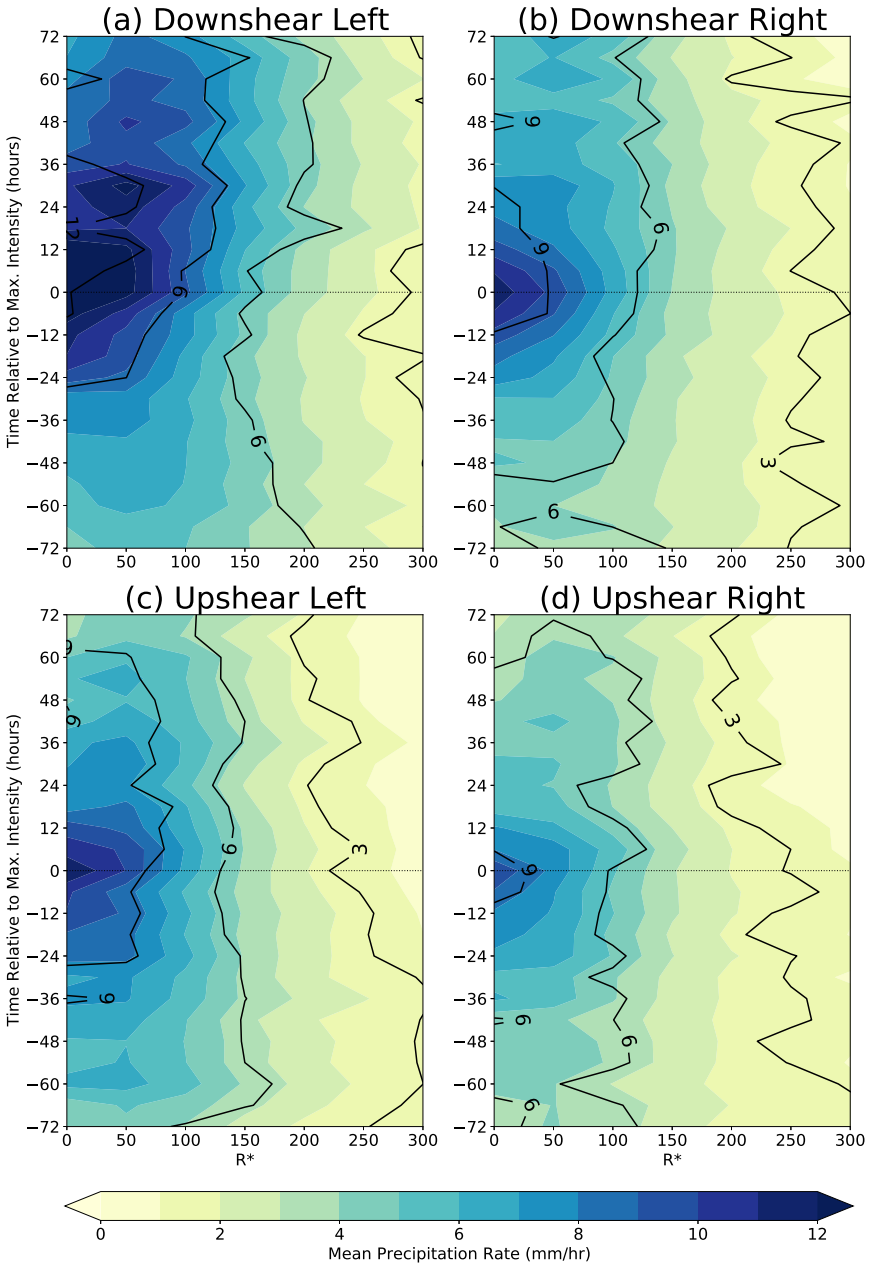

FIG. 2.3. Shear-relative mean GPROF surface precipitation rate in shading, with standard deviation contoured, for the (a) downshear left, (b) downshear right, (c) upshear left, and (d) upshear right quadrants. The x-axes denote the scaled radius while the y-axes denote time in hours relative to the first occurrence of storm lifetime maximum intensity.

convection. Following lifetime maximum intensity, the radial extent of colder TBs continues to be large in the DL quadrant, with a decreasing radial extent in the other quadrants that is

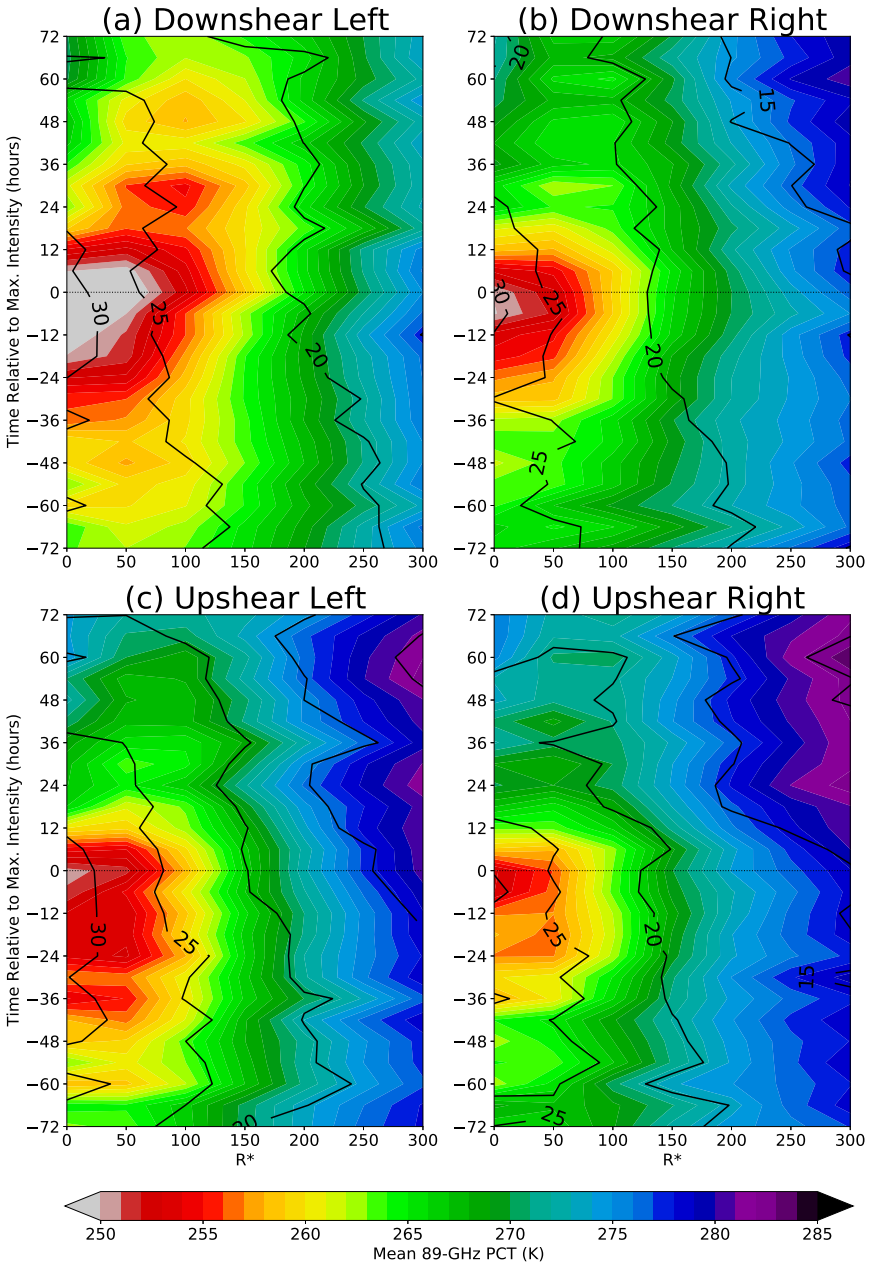

FIG. 2.4. Shear-relative mean 85.5 to 89-GHz brightness temperature in shading, with standard deviation contoured, for the (a) downshear left, (b) downshear right, (c) upshear left, and (d) upshear right quadrants. The x-axes denote the scaled radius while the y-axes denote time in hours relative to the first occurrence of storm lifetime maximum intensity.

most pronounced in the UL and UR quadrants. However, while large precipitation rates lingered after the lifetime maximum intensity, particularly in the DL quadrant (Fig. 2.3), convective intensity denoted by cold TBs did not linger as long, indicating that the large precipitation rates may come from stratiform precipitation.

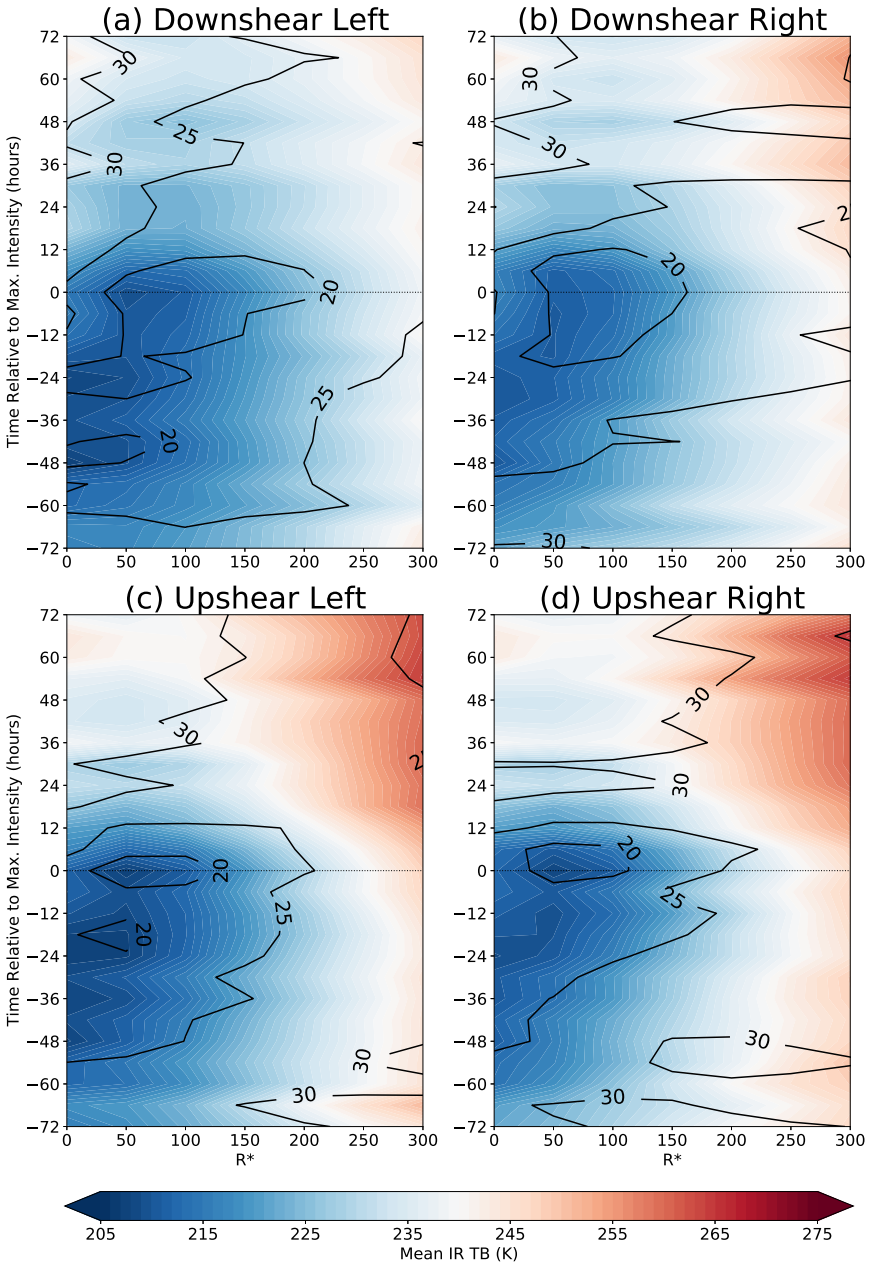

FIG. 2.5. Shear-relative mean infrared brightness temperature in shading, with standard deviation contoured, for the (a) downshear left, (b) downshear right, (c) upshear left, and (d) upshear right quadrants. The x-axes denote the scaled radius while the y-axes denote time in hours relative to the first occurrence of storm lifetime maximum intensity.

The evolution of convection and precipitation are consistent with our current understanding of the tropical cyclone life cycle. Convection (and subsequently precipitation rate) intensifies prior to the lifetime maximum intensity, reaching some semblance of symmetry at lifetime maximum intensity. After the lifetime maximum intensity, convection weakens, presumably

with the weakening of the tropical cyclone intensity, while larger mean precipitation rates linger. The prevalence of convective proxies and precipitation in the downshear quadrants after the lifetime maximum intensity is also indicative of a weakening tropical cyclone due to stronger shear. Generally, the shear-relative distribution of precipitation and convection in the composite analysis agrees with previous studies (e.g. Chen et al. 2006; Hence and Houze 2012), with preferentially larger radial extent and stronger convection and precipitation rate in the downshear quadrants. However, large standard deviations exist in all the analyzed fields because the analysis did not account for the intensity trends prior to and after the first occurrence of the lifetime maximum intensity.

While our analysis is scientifically simple, it is also simple to conduct. We are able to select the basins, sensors, and variables in TC PRIMED with the confident understanding that the data have been inter-calibrated and quality-controlled. In addition, the standardized file format makes ingesting, compositing, and analyzing the huge number of passive microwave and IR observations in TC PRIMED easy. The scientific community should be able to expand on our analysis with more complex approaches, thus advancing our knowledge of tropical cyclones and improving forecasts.

2.5 Conclusions

We have created a long-term, global, research-quality, and open-access dataset of LEO satellite observations of tropical cyclones and associated environmental data called the Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED). TC PRIMED contains tropical cyclone-centric 1) inter-calibrated, multi-channel, multi-imager microwave brightness temperatures, 2) retrieved rainfall from NASA's GPROF, 3) nearly coincident geostationary satellite infrared imagery, and 4) auxiliary data such as tropical cyclone position and intensity, ERA5 fields and derived environmental diagnostics, and satellite precipitation radar variables. TC PRIMED consists of over 168,000 LEO satellite overpasses of 2,101 storms from 1998 to 2019, with user-friendly files that are compliant with two common metadata conventions. We are working to publish TC PRIMED on a more permanent and public online repository.

In this paper, we summarized the data sources we used and the key processing steps we employed to create TC PRIMED. We provided an example of TC PRIMED data from a TRMM overpass of Typhoon Sepat (2007). To demonstrate TC PRIMED's broader utility, we created a shear-relative composite analysis using a select few microwave sensors that are available in TC PRIMED. Our composite analysis corroborates what is already understood about the distribution of convection and precipitation in tropical cyclones. More importantly, our composite analysis demonstrates the huge potential that TC PRIMED presents to the scientific community. By creating TC PRIMED, we have eliminated the tremendous barrier for the scientific community to obtain a large, long-term, global, research-quality, and open-access dataset of LEO satellite observations of tropical cyclones with associated infrared and environmental data — allowing the scientific community to focus solely on research.

CHAPTER 3

PRECIPITATION TYPE CLASSIFICATION IN TROPICAL CYCLONES USING PASSIVE MICROWAVE OBSERVATIONS

3.1 INTRODUCTION

When environmental conditions are favorable, tropical cyclone internal dynamics play a significant role in modulating intensity. One of the key internal dynamical mechanisms that affects intensity is the kinematic and thermodynamic influence of convective and stratiform precipitation on the overall evolution of the tropical cyclone. In order to minimize environmental impacts and investigate the unique role of convective and stratiform precipitation on tropical cyclone intensity, researchers must focus their analysis on observations of tropical cyclones in favorable environments. However, conducting such analysis across a wide number of cases using conventional observational platforms such as ground-based or airborne radar can be challenging due to limited sampling of tropical cyclones in favorable environments.

The Global Precipitation Measurement (GPM; Hou et al. 2014) satellite provides numerous fortuitous observations of tropical cyclones in various environments throughout the globe. The precipitation type classification product from the GPM satellite's dual-frequency precipitation radar (DPR) allows researchers to investigate how the population of convective and stratiform precipitation differs for tropical cyclones undergoing different intensity changes. While the DPR can accurately classify the precipitation type (e.g. convective or stratiform), its narrow swath width (245 km) misses a large portion of the tropical cyclone. Conversely, the GPM Microwave Imager (GMI) has a much wider swath width (885 km), with different microwave channels providing passive radiometric measurements to infer the presence of liquid or ice precipitation. Therefore, to generate a larger sample size of precipitation type observation in tropical cyclones, we train a random forest model to classify the precipitation type with the passive microwave brightness temperatures from the GMI as input predictors and using the DPR classification as reference.

Given the wide-ranging impacts of the kinematic and thermodynamic fields of convective and stratiform precipitation — from the large-scale energy budget to the small-scale storm dynamics — and the enormous amount of satellite passive microwave observations of global precipitating systems, prior research has developed multiple algorithms to classify precipitation type from satellite passive microwave observations. Anagnostou and Kummerow (1997) derived an index to describe the variability of the 85-GHz horizontally-polarized brightness temperature field of the Special Sensor Microwave/Imager (SSM/I). Using ground-based radar observations as reference, they then developed a probability matching algorithm that relates the variability index to the probability of the satellite instantaneous field-of-view (IFOV) containing three different categories of stratiform fraction: 0–40%, 40–70%, and 70–100%. A low variability index indicates higher likelihood that the satellite IFOV contains a large fraction of stratiform precipitation, while a high variability index indicates lower likelihood that the satellite IFOV contains a large fraction of stratiform precipitation.

Expanding on the work of Anagnostou and Kummerow (1997), Hong et al. (1999) used a probability matching technique, multi-channel microwave brightness temperatures, and several brightness temperature variability indices to classify precipitation as convective, stratiform, or mixed for the SSM/I and the Tropical Rainfall Measurement Mission (TRMM) Microwave Imager (TMI) observations. They used radar observations collected during the Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE) and corresponding synthetic microwave data to generate individual algorithms specific for SSM/I and TMI observations over the ocean. Olson et al. (2001) further included polarization information at the 85.5-GHz frequency for the TMI algorithm in Hong et al. (1999). They used the TRMM Precipitation Radar (PR) as their reference, and their algorithm applies to TMI observations over both land and ocean.

Varma and Liu (2010) developed a precipitation type classification algorithm for the TMI using collocated PR data and multi-linear regression. For comparison with their algorithm, Varma and Liu (2010) generated two other outputs using the algorithms of Hong et al. (1999) and Olson et al. (2001), and found that their algorithm performed better. They attributed the

lower performance of the Hong et al. (1999) and Olson et al. (2001) algorithms to the lack of treatment for variable beam filling, where multiple precipitation types exist within the sensor IFOV. While Varma and Liu (2010) meticulously accounted for the elliptical IFOV of the TMI, they noted the difficulty in their own classification due to a different beam filling issue where the majority of their dataset contains TMI data with over 75% no rain within the TMI IFOV. The issues they highlighted demonstrate the challenge in properly representing higher-resolution precipitation radar observations in coarser-resolution passive microwave observations.

Islam et al. (2015) used more complex machine learning methods to classify precipitation type for the TMI. They applied three versions of the adaptive boosting algorithm and two versions of linear discriminant analysis to eighty-five combinations of single channel brightness temperatures, derived brightness temperature quantities (such as the polarization-corrected brightness temperatures), and mathematical combination of brightness temperatures (i.e. adding or subtracting brightness temperatures from two different polarizations or frequencies). They ranked the input predictors based on the most relevant feature combinations first and found that their models performed just as well with a smaller number of important predictor combinations as with a larger number of predictor combinations. Their work highlights the importance of prioritizing the combination of input predictors with the most information.

Using a Bayesian approach, the NASA Goddard Profiling algorithm (GPROF; Kummerow et al. 2015) provides retrieved rainfall and hydrometeor profiles for multiple passive microwave sensors in the GPM constellation. For the GPM sensor, the Bayesian retrieval algorithm estimates the vertical profile of precipitation and surface precipitation rate based on an a-priori database of coincident GPM MI and DPR observations, and ancillary datasets. To create an a-priori database for other sensors in the GPM constellation, the algorithm averages DPR observation to the corresponding microwave IFOV and uses a radiative transfer model. While GPROF does not directly determine if a microwave pixel is convective or stratiform, it provides the convective fraction within the microwave IFOV through the fractional contribution of convective surface precipitation rate to the total surface precipitation rate. The convective fraction is dependent on the IFOV of the sensor. Since we use only the higher-resolution GMI

data for our algorithm, we compare our results with the GPROF convective fraction variable for the GMI.

More recently, Petković et al. (2019) trained a deep neural network algorithm to classify convective and stratiform precipitation for the GMI. They used a combined radar and radiometric algorithm as their reference. However, the DPR exclusively determines the fraction of convective precipitation within the GMI IFOV. Their input predictors are of 3×5 brightness temperature pixel batches centered on the input class from all 13 GMI frequencies. Petković et al. (2019) used the 3×5 brightness temperature pixel batches to avoid manually calculating texture information, like the Hong et al. (1999) and Olson et al. (2001) algorithms. Petković et al. (2019) reduced the precipitation rate bias in GPROF by generating two sets of a-priori database for GPROF — one for convective precipitation and one for stratiform precipitation.

While most of the authors cited above developed their algorithms for the TMI, we are interested in using the higher resolution observations from the GMI. For example, the IFOV for the TMI 10.65-GHz frequency is approximately 41×68 km, whereas the IFOV for the GMI 10.65-GHz frequency is approximately 19×32 km. And while Petković et al. (2019) used GMI observations in their approach, their precipitation type field is too coarse. We seek to instead obtain a higher resolution precipitation type field using the finer IFOV of the GMI higher frequencies (e.g., 4.4×7.2 km at 89.0 GHz). In addition, as we will show later, GPROF overestimates the convective fraction in regions with little to no precipitation, which will bias our subsequent analysis on the distribution of convection in tropical cyclones.

Finally, while a more generalized precipitation type classification algorithm like that from Petković et al. (2019) is useful for a variety of precipitating systems, it may not necessarily be optimal for tropical cyclone studies alone. Brown et al. (2016) showed that the strong horizontal flow in tropical cyclones produce a precipitation profile that is different from other precipitating systems. Therefore, they developed a hurricane version of GPROF called HGPROF, which improved the rainfall retrievals in tropical cyclones. Since HGPROF uses the same Bayesian approach as GPROF, it suffers from the same bias in convective fraction. For these reasons, we train our own precipitation type classification algorithm for the GMI with

the goal of using the algorithm to understand how the population of convective and stratiform precipitation vary in tropical cyclones undergoing different intensity changes. We seek to not only develop our own algorithm, but to also understand reasons behind our algorithm's successes and failures.

For our algorithm, we use the random forest classifier. Our input predictors are GPM observations of tropical cyclones, which we obtain from the Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED; Razin et al. 2021). TC PRIMED is a dataset of satellite passive microwave observations centered on tropical cyclones and includes ancillary data such as infrared brightness temperatures and re-analysis fields. However, we use only the GPM MI and DPR data to reduce dependency on other data sources. In section 2, we discuss data preparation, and model selection and training. We discuss our model's performance in section 3 and key characteristics of our model in section 4. In section 5, we offer our concluding remarks.

3.2 METHODS

3.2.1 GMI and DPR

For our precipitation type classification algorithm, we use observations from the GMI and the DPR. Our input predictors are brightness temperature values and derived quantities from the GMI. The GMI observes terrestrial radiation at the following frequencies: 10.65, 18.7, 23.8, 36.64, 89.0, 166.0, 183.31±3.0, and 183.31±7.0 GHz. The GMI divides observations at each frequency into a vertical and horizontal polarization component except for the 23.8, 183.31±3.0, and 183.31±7.0 GHz frequencies, at which only the the vertical polarization component exists. The IFOV for each channel ranges from an ellipse approximately 19.4 × 32.1 km for the 10.65-GHz frequency to an ellipse approximately 3.8 × 5.8 km for the 183.31-GHz frequency. To avoid dealing with the complexity of different land surface emissivities, we focus our algorithm to classify precipitation type only for pixels over the ocean.

We train the GMI brightness temperatures and derived quantities on the precipitation type classification from the GPM DPR. The GPM DPR actively senses precipitation at two different frequencies, in the Ka-band (35.5-GHz) and the Ku-band (13.6-GHz). The Ka- band is more

sensitive to attenuation and initially had a narrower swath width of 120 km compared to 245 km swath width of the Ku-band. While the Japanese Aerospace Development Agency (JAXA), the agency that manages the DPR data, expanded the swath width of the Ka-band observation in May of 2018, we use observations from the Ku-band for consistency throughout the period of our training dataset. Readers can find an extensive discussion on the precipitation type classification algorithm for the GPM DPR in Awaka et al. (2016) and Iguchi et al. (2018). For completeness, we summarize the algorithm here.

The algorithm first looks for the presence of a bright band in the vertical radar reflectivity profile. If a bright band exists, then the algorithm classifies the pixel as stratiform. If a bright band exists but the attenuation-corrected radar reflectivity beneath the bright band exceeds 46 dBZ, then the pixel is convective. In addition, if the attenuation-corrected radar reflectivity beneath the bright band exceeds the maximum reflectivity value in the bright band, the pixel is also convective. If a bright band does not exist, the pixel is convective if the radar reflectivity exceeds 40 dBZ at any altitude below the storm top. Finally, the pixel and its surrounding pixels are convective if the radar reflectivity of the central pixel exceeds a certain reflectivity threshold relative to the mean reflectivity of its surroundings, similar to the algorithm of Steiner et al. (1995). If the algorithm detects precipitation at altitudes higher than the analyzed 0°C level, but could not detect precipitation in the lower levels, the algorithm classifies the pixel as other. Iguchi et al. (2018) hypothesized that this other category is mere clouds without precipitation, or noise. Based on their description, we surmised that the other precipitation type resembles overhanging precipitation or cirrus. While this "other" precipitation type rarely occurs in our dataset, we observed cases where this precipitation type does occur in a few GPM DPR observations of a tropical cyclone. Therefore, we include this precipitation type to avoid falsely classifying them as either convective or stratiform.

Finally, since we use TC PRIMED as our data source of GPM observations, our dataset also contains a shallow rain type. Razin et al. (2021) pointed out that the original GPM DPR precipitation type classification algorithm groups shallow rain with deeper convection. Shallow rain is any precipitating pixel where the storm top height is at most 1 km below the analyzed

0°C level. Razin et al. (2021) reasoned that the passive microwave signatures of shallow rain and deeper convection can be vastly different, and therefore chose to further separate shallow rain from the convective rain type. Overall, we attempt to classify five precipitation types: no rain, stratiform, convective, other, and shallow rain. However, given that other and shallow rain are not as numerous and do not have a significant impact on a tropical cyclone as do convective and stratiform rain types, we gauge the success of our algorithm based on its ability to correctly classify convective and stratiform precipitation.

3.2.2 Re-gridding Information onto a Common Grid

In order to have a consistent training dataset, we must first re-grid our input data onto a common grid. The GMI observes brightness temperatures on two sets of grids or swaths. The first swath (S1) contains brightness temperature observations for the 10.65-GHz to the 89.0-GHz frequencies. The second swath (S2) is slightly narrower than S1 and contains brightness temperature observations for the 166.0-GHz and 183.31-GHz frequencies. Figure 3.1a shows an example of the location of brightness temperature observations for S1 and S2. We collocate the brightness temperature observations from S2 to S1 using the following equation:

$$\mathrm{TB}_{\mathrm{S1}} = \frac{\sum w \, \mathrm{TB}_{\mathrm{S2}}}{\sum w} \tag{3.1}$$

$\mathrm{TB}_{\mathrm{S1}}$ is the S2 brightness temperature collocated onto S1, $\mathrm{TB}_{\mathrm{S2}}$ is all of the brightness temperature observations in S2, and w is a weighting function, where:

$$w = \exp\left(-\frac{r^2}{\mathrm{FOV}}\right) \tag{3.2}$$

Here, r is the distance from a particular S1 grid point to all of the S2 grid points. FOV is a field-of-view parameter, which we use to control the smoothness of the weighting function. When re-gridding S2 brightness temperatures to S1, we select FOV = 5 km, which is the approximate IFOV for the S2 brightness temperature observations. We note that our re-gridding method is purely mathematical, and that the proper re-gridding method requires detailed considerations of the sensor such as the elliptical shape of the IFOV and the antenna gain function (Backus and Gilbert 1970). For computing efficiency, we opted for the simpler

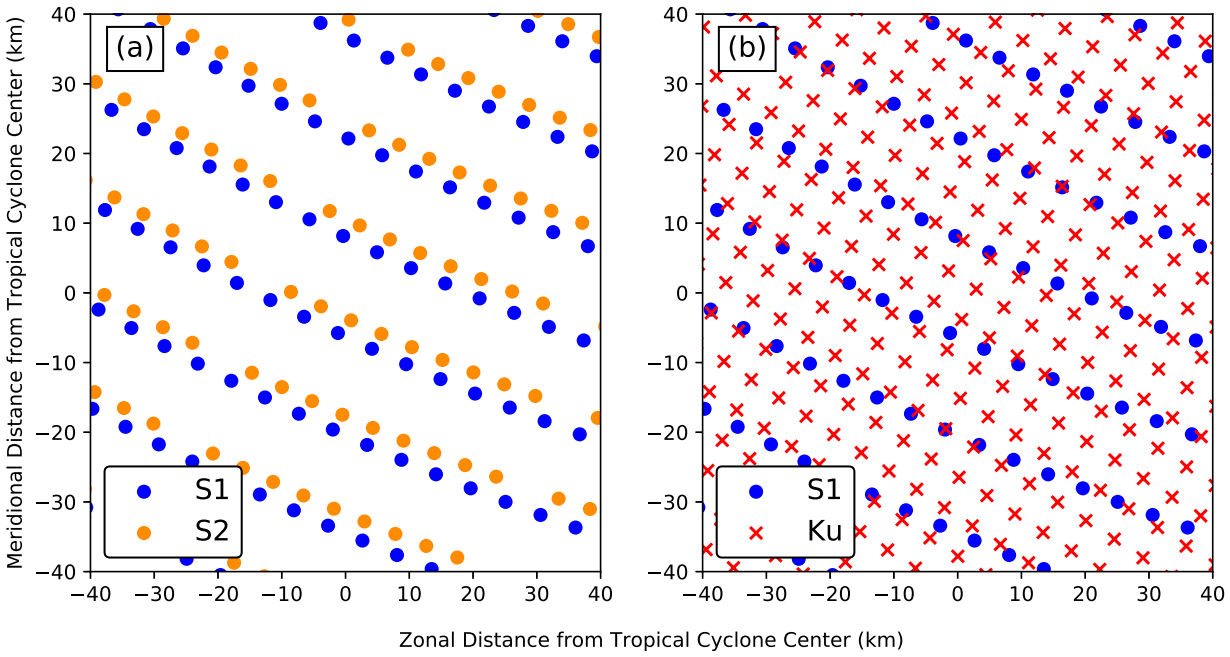

FIG. 3.1. Relative location of (a) the S2 observations to the S1 observations and (b) the DPR observations to the S1 observations. The IFOV of the observations is not to represented.

mathematical re-gridding method. Finally, since S2 is narrower than S1, we make sure not to extrapolate observations from S2 to S1 and only re-grid the S2 brightness temperature values to S1 if the corresponding S1 grid point is within S2. Figure 3.2 shows the 183.31±3.0-GHz brightness temperature observations at its original S2 grid points and its re-gridded field on S1 grid points. The re-gridding process using FOV = 5 km sufficiently captures the brightness temperature observations on the original swath.

To re-grid the GPM DPR observations to S1, we calculate the fractional weight of each precipitation type and select the precipitation type with the highest fractional weight. In mathematical form:

$$\mathrm{PT}_{\mathrm{S1}} = \max\left(\frac{w_i}{\sum w_i}\right) \tag{3.3}$$

$\mathrm{PT}_{\mathrm{S1}}$ is the precipitation type from the GPM DPR re-gridded to S1, and i = no rain, stratiform, convective, other, shallow rain. We adopt this method to represent the most dominant precipitation type in the GMI IFOV and thus, simplifying our approach. For the weighting

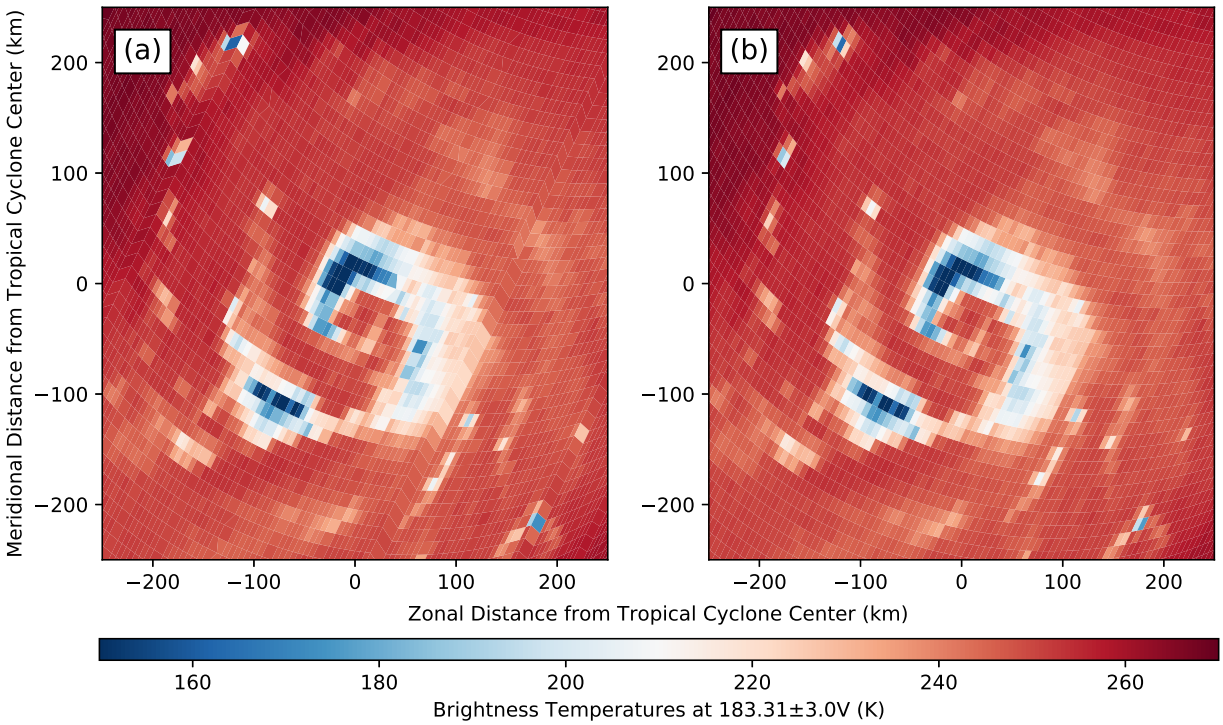

FIG. 3.2. Brightness temperature observations in the 183.31±3.0-GHz channel (a) on its original swath (S2) and (b) interpolated onto S1.

function (eq. 1.2), we use FOV = 5 km, which, as we mentioned, is the approximate IFOV of the highest resolution GMI observations as well as the approximate IFOV of the GPM DPR pixel. We chose FOV = 5 km to preserve the details in the precipitation field as much as possible. Using a larger FOV will lead to a smoother precipitation field. To test the sensitivity of our model to the choice of FOV, we generate a separate precipitation type field where we re-grid the DPR precipitation type to S1 using FOV = 25 km in the weighting function. Figure 3.1b shows an example of the location of DPR observations relative to location of S1 observations, while Fig. 3.3 shows the original precipitation type classification field from the DPR and the re-gridded precipitation type classification on S1 using FOV = 5 km and FOV = 25 km.

3.2.3 Preparing the Input Variables

After re-gridding all of the input onto a common grid, we further include derived variables as additional input predictors on top of the multi-frequency passive microwave observations. We include polarization-corrected brightness temperature (PCT) for two frequencies. PCTs

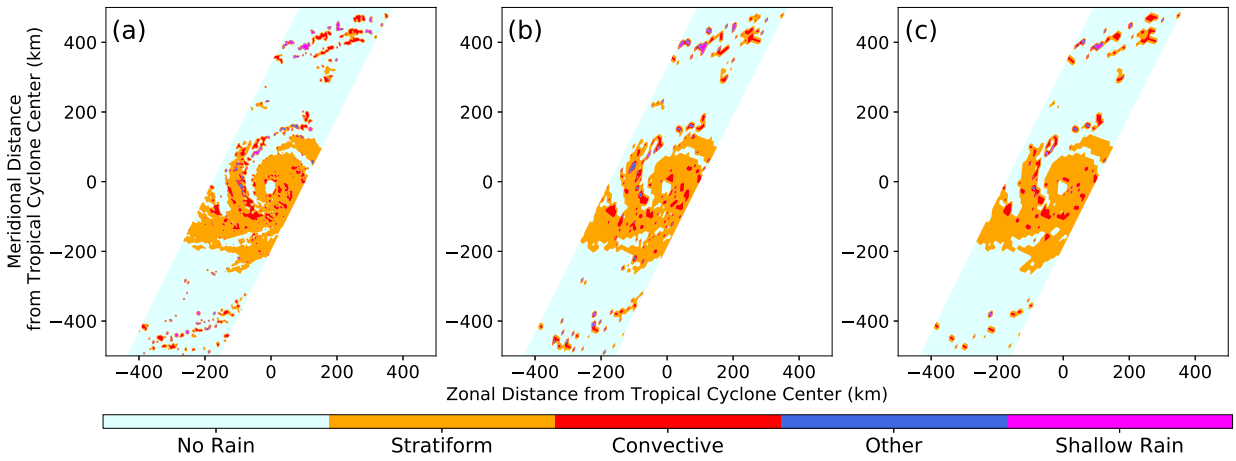

FIG. 3.3. (a) The raw precipitation type observation from the DPR, (b) the precipitation type observation from the DPR interpolated onto S1 using FOV = 5 km in the weighting function, and (c) the precipitation type observation from the DPR interpolated onto S1 using FOV = 25 km in the weighting function.

use observations from the horizontal and vertical polarizations of a particular frequency to remove ambiguity in the brightness temperatures due to varying surface emissivities. PCT observations are sensitive to scattering of terrestrial radiation by large precipitation-sized ice particles. We calculate the PCT for the 36.64-GHz and 89.0-GHz frequencies using the equations derived and coefficients suggested in Cecil and Chronis (2018). The equations are:

$$36.64\text{PCT} = [2.15 \times \text{TB}(36.64\text{V})] - [1.15 \times \text{TB}(36.64\text{H})]$$

$$89.0\text{PCT} = [1.7 \times \text{TB}(89.0\text{V})] - [0.7 \times \text{TB}(89.0\text{H})]$$

Here, TB(36.64V) and TB(36.64H) are the respective vertical and horizontal polarization brightness temperature observations at 36.64-GHz, and TB(89.0V) and TB(89.0H) are the respective vertical and horizontal polarization brightness temperature observations at 89.0-GHz.

At around 37-GHz, the ocean surface is the main contributor to the polarization difference in brightness temperatures (Petty 1994). As precipitation increasingly occupies the 37-GHz IFOV, the difference in the vertical and horizontal polarizations decreases as the sensor increasingly receives emission from liquid precipitation. Therefore, to help in delineating

raining and non-raining passive microwave pixels, we include the 36.64-GHz polarization difference, which we refer to as the 36.64-GHz emission index, given by:

$$\mathrm{EI}_{36.64} = \mathrm{TB}(36.64\mathrm{V}) - \mathrm{TB}(36.64\mathrm{H})$$

In contrast to the horizontally homogeneous stratiform precipitation, the strong and localized vertical velocities in convective precipitation produce more discrete and cellular regions of strong precipitation and large ice particles aloft. Therefore, convective precipitation have a more variable passive microwave representation, particularly in the ice-scattering frequencies. For example, an 89.0-GHz passive microwave pixel over convection will have a much stronger ice-scattering signature relative to the surrounding regions, producing a more peaked or variable brightness temperature field compared to the brightness temperature field of non-raining or stratiform precipitation. This physical concept is the reason behind the introduction of the variability indices in Anagnostou and Kummerow (1997) and Hong et al. (1999). To add some texture information to our model, we use the three variability indices derived in Anagnostou and Kummerow (1997) and Hong et al. (1999), which they referred to as VM, VI, and VC. These equations describe texture information based on mathematical relationships between the central passive microwave pixel and the surrounding eight pixels.

VM describes the maximum brightness temperature enhancements due to liquid water emission in the liquid water emission channels (e.g. 36.64-GHz) and the maximum brightness temperature depression due to ice scattering in the ice-scattering channels (e.g. 89.0-GHz). We calculate VM for the 36.64-GHz vertical polarization channel to exploit the larger dynamic range of the vertical polarization and the 89.0-GHz PCT to avoid ambiguities from varying surface emissivities in either the vertical or horizontal polarization observations at 89.0-GHz. With TB_c being the brightness temperature value of the center pixel and TB_i being the brightness temperature values for the surrounding eight pixels, the VM calculation for the vertical polarization 36.64-GHz and 89.0-GHz PCT are:

$$\mathrm{VM}_{36.64\mathrm{V}} = \max[\mathrm{TB}_c(36.64\mathrm{V}) - \mathrm{TB}_i(36.64\mathrm{V})] \geq 0$$

$$\mathrm{VM}_{89.0\mathrm{PCT}} = \min[\mathrm{TB}_c(89.0\mathrm{PCT}) - \mathrm{TB}_i(89.0\mathrm{PCT})] \leq 0$$

To focus on the variability due to ice scattering in convection, we use only the 89.0-GHz PCT for VC and VI, given as:

$$\mathrm{VC}_{89.0\mathrm{PCT}} = \mathrm{TB}_c - \frac{1}{8}\sum_{i=1}^{8}\mathrm{TB}_i(89.0\mathrm{PCT})$$

$$\mathrm{VI}_{89.0\mathrm{PCT}} = \frac{1}{8}\sum_{i=1}^{8}|\mathrm{TB}_c(89.0\mathrm{PCT}) - \mathrm{TB}_i(89.0\mathrm{PCT})|$$

Table 3.1 shows the list of input predictors we use in our model. For our model training, we use all of the GPM observations in TC PRIMED from the 2014 to the 2019 tropical cyclone season. However, we omit the 2017 tropical cyclone season to analyze the applicability of the model's prediction. In total, our input dataset contains over 14.3 million data points.

3.2.4 Model Setup and Training

For our model, we use the random forest from the Scikit-Learn package (Pedregosa et al. 2011) version 0.23.2 for Python. The Python version is 3.8.5 (Van Rossum and Drake 2009). We choose random forest for its relative simplicity and representation of class probabilities based on the number of trees that classify that class. We train our model on 80% of our input data and test its performance on the remaining 20%. We randomize this train/test split with a random seed of 42 and stratify them to ensure the same fraction of precipitation types in each split. Before we train the model to classify precipitation, we first search for the best model hyperparameters using GridSearchCV. We set the number of trees in the forest to 50, and search for the best model depth and minimum number of samples in a leaf node. The range of model depths is: 10, 20, 30, 40, 50, 60, 70; while the range of minimum number of samples in a leaf node is: 5, 15, 25, 35, 45. We set the rest of the model hyperparameters to their default values

TABLE 3.1. List of input predictors used to train our model. We calculate derived quantities from their corresponding raw microwave brightness temperatures.

Raw Brightness Temperature	Derived Quantities
10.65-GHz V & H	
18.7-GHz V & H	
23.8-GHz V	
36.64-GHz V & H	$EI_{36.64}$, 36.64PCT & $VM_{36.64V}$
89.0-GHz V & H	89.0PCT, $VM_{89.0PCT}$, $VC_{89.0PCT}$, & $VI_{89.0PCT}$
166.0-GHz V & H	
183.31±3.0-GHz V	
183.31±7.0-GHz V	

with the exception of employing a balanced class weighting to account for the class imbalance in our input data. To avoid over-fitting, we separate our training dataset into 5 folds for cross-validation, with each fold stratified to have a similar number of each class. Finally, we use the macro-averaged F1 score as our scoring metric.

The F1 score considers both the model's recall and precision in its predictions. Recall measures the number of correct model predictions for a particular class relative to the total number of samples of that class in the dataset. Precision measures how often the model correctly predicts a class every time the model predicts that class. In mathematical form:

$$\text{recall} = \frac{TP}{TP+FN}$$
$$\text{precision} = \frac{TP}{TP+FP}$$

Here, TP is the number of true positives, FN is the number of false negatives, and FP is the number of false positives in the model's predictions. The F1 score is thus:

$$\text{F1} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}}$$

The F1 score ranges between 0 and 1, with a higher F1 score indicating a better model. For a multiclass problem, GridSearchCV calculates the F1 score for each individual class. With our macro-averaged F1 scoring metric, GridSearchCV searches for the model with the best simple

average of the F1 score of each class. In other words, the F1 score for each class has a weight of one. The best macro-averaged F1 score from the GridSearchCV step is 0.5857, with the best model depth and minimum number of samples in a leaf node being 50 and 5, respectively.

After we obtain the best model hyperparameters, we employ a calibration step in our model training. A model is calibrated when the likelihood of a correct prediction is proportional to the probability that the model assigned to its prediction. For example, a predicted class probability of 75% should be correct 75% of the time. Random forest models are not calibrated and may be over-confident or under-confident in making a classification based on its uncalibrated probability. The calibration step cannot completely calibrate the model due to random noise in the dataset. Rather, the calibration step brings the model closer to being calibrated, but may lead to a slight decrease in the F1 score. Using the model hyperparameters we obtain through GridSearchCV, we train and calibrate our model with cross-validation. The model parameters are 50 trees, a depth of 50, and a minimum number of samples in a leaf of 5. The class weights are balanced and we keep the rest of the parameters to their default values. Identical to the GridSearchCV step, we separate our training data into 5 folds, with each fold stratified to have similar number of each class. The calibration step involves fitting the model on four folds and calibrating on the fifth fold. The macro-averaged F1 score for the calibrated model is 0.5730.

3.3 Results

We first present the results of the model calibration step using a reliability diagram (Fig. 3.4). A reliability diagram presents the probability of observing a true class given the model prediction probability. A model with perfect reliability has the probability of observing a true class that corresponds to the model prediction probability (the 1:1 line on a reliability diagram). Using the input predictors from the testing dataset, we obtain the uncalibrated and calibrated model prediction probability for no rain (Fig. 3.4a), stratiform (Fig. 3.4b), and convective (Fig. 3.4c) precipitation. We use our truth labels from the testing dataset and collapse the labels into a binary problem where we assign a value of 1 to the precipitation type of interest and a value of 0 to the other precipitation types — which we do separately for the no rain, stratiform, and convective precipitation type. We then compute the reliability curve for each

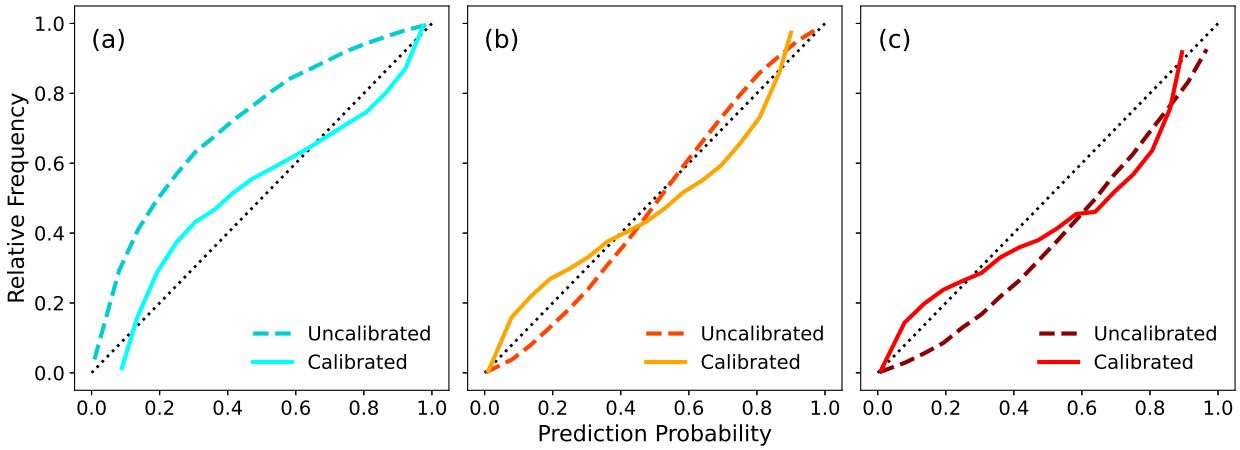

FIG. 3.4. Reliability diagram for (a) no rain, (b) stratiform, and (c) convective precipitation type. Darker dashed lines in each panel show the reliability curve for the uncalibrated model, while the lighter solid lines show the reliability curve for the calibrated model.

precipitation type using Scikit-Learn's reliability curve package, using 18 probability bins of equal bin width.

Figure 3.4a shows that the uncalibrated model was under-confident in its predictions of no rain. For example, whenever the model predicts 40% probability of no rain (prediction probability of 0.4), the number of no rain pixels that actually exist in this prediction probability category is over 60% (relative frequency of 0.6). The calibrated model produces a reliability curve that is closer to the 1:1 line, with still some under-confidence in the lower prediction probability range (below 0.7). For stratiform precipitation, the reliability curves of both the uncalibrated and calibrated models are close to the 1:1 line (Fig. 3.4b). But the reliability curve of the calibrated model is slightly further away from the 1:1 line. The reduction in the reliability of the calibrated model in predicting stratiform precipitation is not tied to the improvements in the reliability of the calibrated model in predicting no rain, as Scikit-Learn's calibration routine calibrates the probability of each precipitation type separately. For convective precipitation, the calibrated model is closer to the 1:1 line in the low range of prediction probability (below 0.3; Fig. 3.4c). At the higher range of prediction probability (between 0.6 and 0.9), the calibrated model is more over-confident than the uncalibrated model. The improvement in model reliability, particularly for the no rain classification, justifies the calibration step that we employed. However, the reasons behind the slight reduction in model

TABLE 3.2. Confusion matrix for the uncalibrated model. The top number in each panel corresponds to the absolute number of cases that belong in that panel, with their row and column totals located in the far right column and bottom row, respectively. The top percentages in each panel is the recall and adds up to 100% in each row. The bottom and bold percentages in each panel is the precision and adds up to 100% in each column. The shaded cells highlight instances when our model correctly classifies a precipitation type.

		Predicted					
		No Rain	Stratiform	Convective	Other	Shallow	
Truth	No Rain	2,271,865 (95.7%) **(97.8%)**	43,042 (1.8%) **(12.3%)**	33,854 (1.4%) **(24.4%)**	13,963 (0.6%) **(44.2%)**	10,364 (0.4%) **(53.0%)**	2,373,088
	Stratiform	15,515 (4.7%) **(0.7%)**	272,460 (82.4%) **(77.7%)**	35,437 (10.7%) **(25.6%)**	6,809 (2.1%) **(21.5%)**	518 (0.2%) **(2.6%)**	330,739
	Convective	12,345 (11.3%) **(0.5%)**	27,943 (25.6%) **(8.0%)**	65,697 (60.3%) **(47.4%)**	431 (0.4%) **(1.4%)**	2,609 (2.4%) **(13.3%)**	109,025
	Other	10,045 (35.7%) **(0.4%)**	7,006 (24.9%) **(2.0%)**	684 (2.4%) **(0.5%)**	10,382 (36.9%) **(32.9%)**	4 (0%) **(0%)**	28,121
	Shallow	13,162 (59.4%) **(0.6%)**	113 (0.5%) **(0%)**	2,808 (12.7%) **(2.0%)**	13 (0.1%) **(0%)**	6,068 (27.4%) **(31.0%)**	22,164
		2,322,932	350,564	138,480	31,598	19,563	

reliability for stratiform and convective classification is unclear and is subject to future analyses.

We now investigate the difference in prediction skill between the uncalibrated and the calibrated models. Tables 3.2 and 3.3 show the confusion matrix for the uncalibrated and calibrated model, respectively. The calibrated model has a higher recall for non-raining pixels compared to the uncalibrated model (97.7% to 95.7%), but lower precision (96.7% to 97.8%). The reverse is true for stratiform and convective precipitation. For stratiform precipitation, the calibrated model has a marginally lower recall for stratiform pixels compared to the uncalibrated model (81.9% to 82.4%). However, the calibrated model has a higher precision whenever it predicts the stratiform precipitation type (80.7% to 77.7%). Similarly, the calibrated model has a lower recall for convective pixels than the uncalibrated model (52.2% to 60.3%), but a higher precision (54.7% to 47.4%).

TABLE 3.3. Confusion matrix for the calibrated model. The top number in each panel corresponds to the absolute number of cases that belong in that panel, with their row and column totals located in the far right column and bottom row, respectively. The top percentages in each panel is the recall and adds up to 100% in each row. The bottom and bold percentages in each panel is the precision and adds up to 100% in each column. The shaded cells highlight instances when our model correctly classifies a precipitation type.

		Predicted					
		No Rain	Stratiform	Convective	Other	Shallow	
Truth	No Rain	2,319,344 (97.7%) **(96.7%)**	29,295 (1.2%) **(8.7%)**	16,492 (0.7%) **(15.9%)**	4,621 (0.2%) **(29.5%)**	3,336 (0.1%) **(38.1%)**	2,373,088
	Stratiform	26,956 (8.2%) **(1.1%)**	270,880 (81.9%) **(80.7%)**	28,570 (8.6%) **(27.5%)**	4,019 (1.2%) **(25.7%)**	314 (0.1%) **(3.6%)**	330,739
	Convective	21,282 (19.5%) **(0.9%)**	29,226 (26.8%) **(8.7%)**	56,859 (52.2%) **(54.7%)**	106 (0.1%) **(0.7%)**	1,552 (1.4%) **(17.7%)**	109,025
	Other	14,610 (52.0%) **(0.6%)**	6,253 (22.2%) **(1.9%)**	341 (1.2%) **(0.3%)**	6,917 (24.6%) **(44.2%)**	0 (0%) **(0%)**	28,121
	Shallow	16,860 (76.1%) **(0.7%)**	42 (0.2%) **(0%)**	1,705 (7.7%) **(1.6%)**	2 (0%) **(0%)**	3,555 (16.0%) **(40.6%)**	22,164
		2,399,052	335,696	103,967	15,665	8,757	

Relative to the uncalibrated model, the calibrated model classifies more stratiform and convective precipitation type as no rain (8.2% of stratiform and 19.5% of convective pixels classified as no rain). However, whenever the calibrated model predicts stratiform or convective precipitation type, its classifies fewer no rain pixels as either rain type (lower false alarm rate with respect to stratiform or convective precipitation type). Hereafter, we discuss our results based on the performance of the calibrated model. We choose the calibrated model because of its lower false alarm rates for stratiform and convective precipitation. The lower false alarm rates for stratiform and convective precipitation signify higher confidence that a stratiform or convective classification by our model is not based on "artificial" data.

Our model correctly classifies 90.5% of stratiform precipitation and 79% of convective precipitation as either stratiform or convective. Our model incorrectly classifies 1.5% or fewer of

either stratiform or convective type as other and shallow combined. Therefore, while delineating convective and stratiform precipitation requires further improvement, we can be certain of our model's ability to differentiate significant precipitation like stratiform and convective precipitation from no rain.

3.4 Discussion

3.4.1 Understanding Failure Modes

To further evaluate our model's performance, we present a couple of examples from a GPM observation of Hurricane Irma (2017) and Typhoon Talim (2017) in Figures 3.5–3.8. Recall that we left out the 2017 tropical cyclone season from model training and testing. Looking first at Hurricane Irma, we see that our model does well in capturing the general structure of the storm (Fig. 3.5f). Our model captures i) the convective ring of the eyewall, ii) the non-raining region in the eye, ii) the overall stratiform rain surrounding the eye, iv) the other precipitation type and convective band to the south-southwest of the eye, and v) the rainband to the north-northwest of the eye. Our model does not predict any shallow rain in this view even though some shallow rain exist in the observations (Fig. 3.5d,e). More importantly, relative to GPROF (Fig. 3.5c), our model does well in separating non-raining from precipitating regions. For example, GPROF predicts really high convective fraction to the west and northwest of the storm (Fig. 3.5c), where the observation and model show no rain (Fig. 3.5d,e,f).

The high convective fraction from GPROF is the result of the algorithm's Bayesian averaging scheme, where even though the algorithm is fairly certain that a pixel is non-raining, there is still some contribution from the raining pixels in its a-priori database to the final solution. Figure 3.6 shows the locations where our model incorrectly classifies no rain, stratiform, and convective precipitation type. Our model's confusion between no rain and stratiform pixels (e.g. predicts no rain for a stratiform pixel or predicts stratiform for a no rain pixel) occurs on the edges of the two precipitation types (Fig. 3.6c,d). The same is true for our model's confusion between no rain and convective pixels, with the addition that our model also struggles in classifying the more discrete convective precipitation to the northwest of the eye (Fig. 3.6e,f). In delineating stratiform and convective precipitation, our model similarly struggles

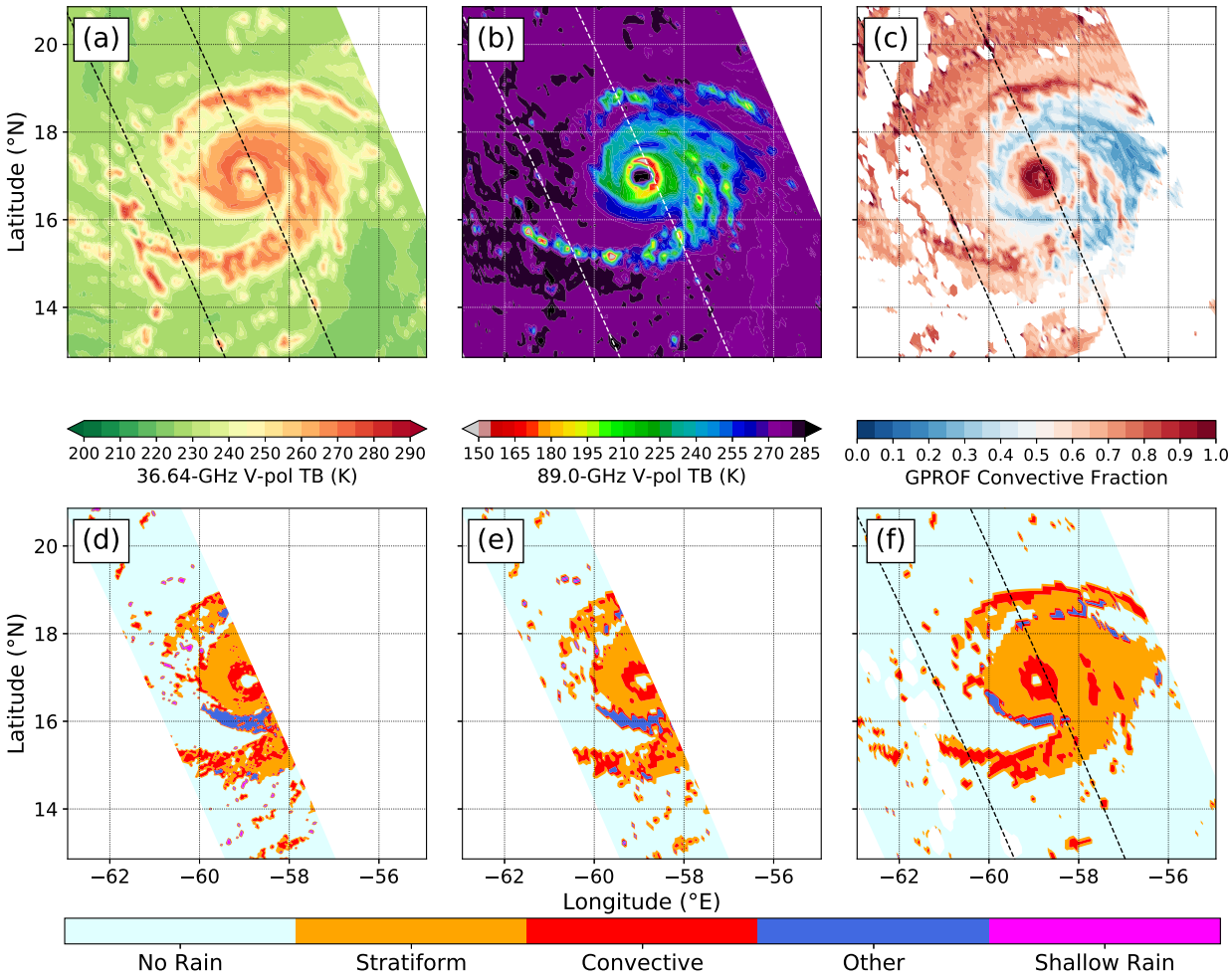

FIG. 3.5. GPM observation of Hurricane Irma (2017) at 16:52:05 UTC on September 5th, 2017 showing (a) 36.64-GHz vertical polarization brightness temperatures, (b) 89.0-GHz vertical polarization brightness temperatures, (c) GPROF convective fraction output, (d) precipitation type classification from the DPR, (e) precipitation type classification re-gridded onto S1, and (f) precipitation type classification from our random forest model.

on the edge of the two precipitation types. In addition, our model also struggles to delineate the precipitation type when the precipitation field is disorganized or random, such as within the rainband to the south-southwest of the eye (Fig. 3.6a,g,h).

The example from Hurricane Irma shows that our model struggles to classify precipitation types on the edges of two precipitation types, when the precipitation field is discrete, and/or when the precipitation field is more randomly distributed. The listed difficulties in accurate model classification point towards an issue with non-uniform beam filling, where two or more precipitation types exist within the satellite IFOV. In addition, parallax error can also

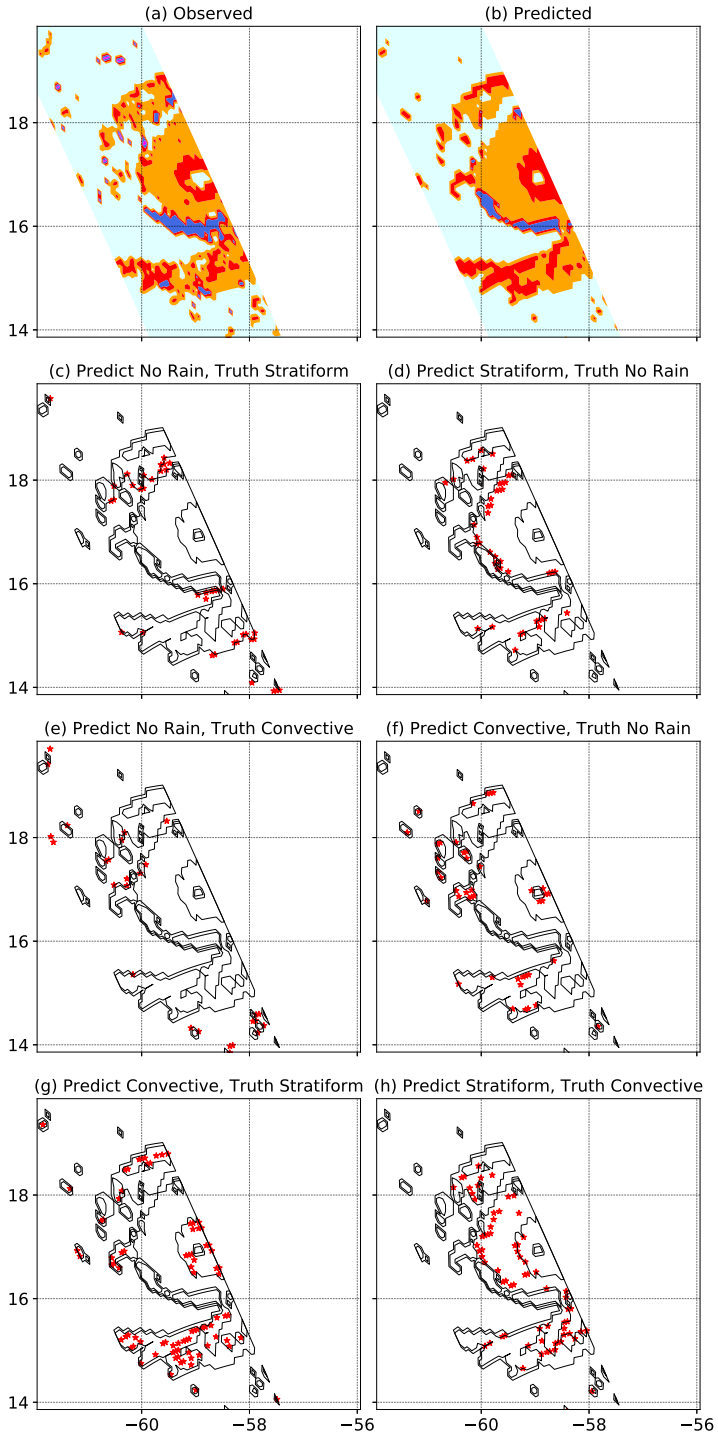

FIG. 3.6. Hurricane Irma's (a) DPR precipitation type classification re-gridded onto S1 as in Fig. 3.5e, (b) precipitation type classification from our random forest model as in Fig. 3.5f, and locations as marked by red stars of when our model (c) predicts no rain for a stratiform pixel, (d) predicts stratiform for a no rain pixel, (e) predicts no rain for a convective pixel, (f) predicts convective for a no rain pixel, (g) predicts convective for a stratiform pixel, and (h) predicts stratiform for a convective pixel.

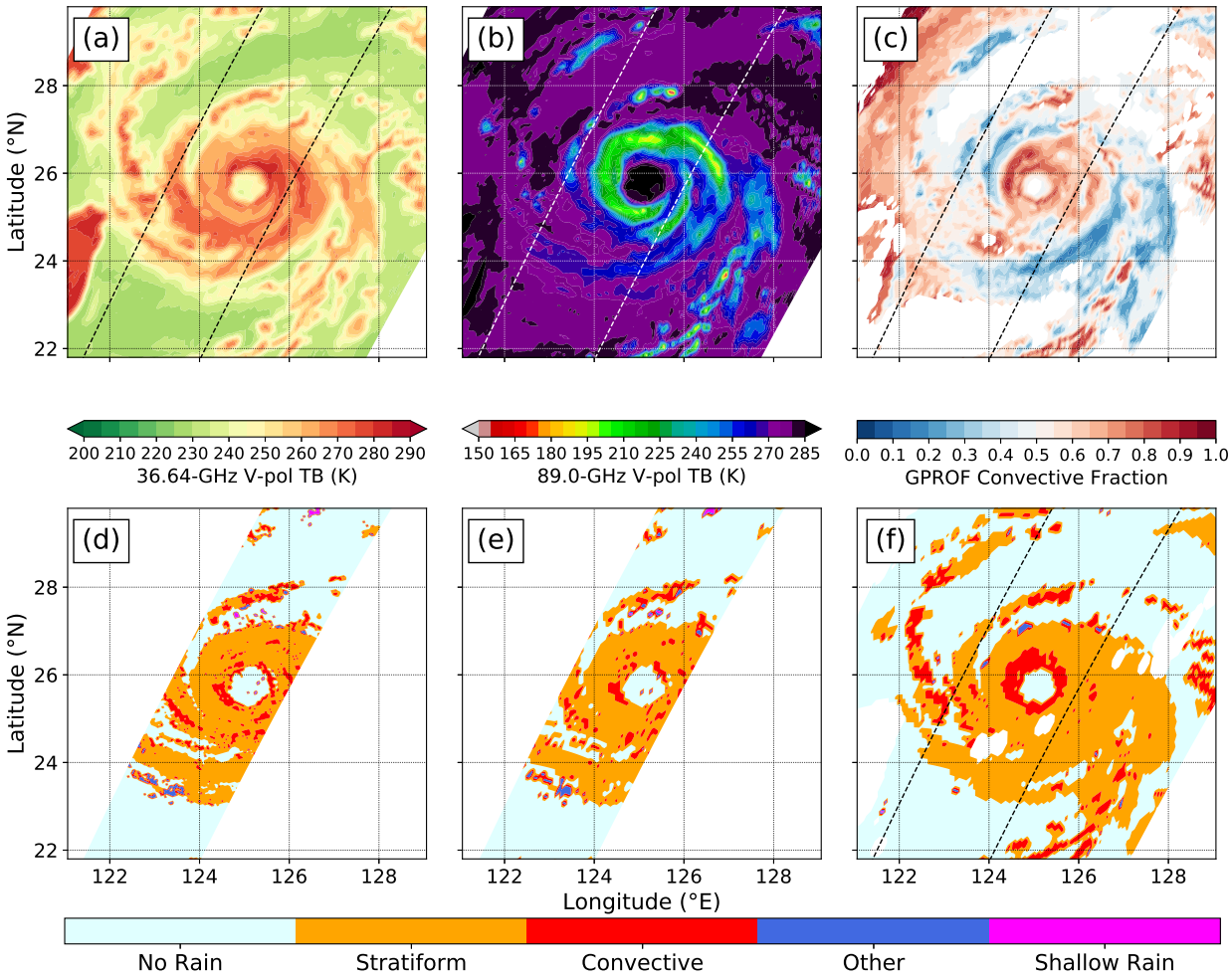

FIG. 3.7. GPM observation of Typhoon Talim (2017) at 15:36:15 UTC on September 13th, 2017 showing (a) 36.64-GHz vertical polarization brightness temperatures, (b) 89.0-GHz vertical polarization brightness temperatures, (c) GPROF convective fraction output, (d) precipitation type classification from the DPR, (e) precipitation type classification re-gridded onto S1, and (f) precipitation type classification from our random forest model.

contribute to more noise in training our model, since the DPR senses precipitation at nadir while the GMI has an off-nadir viewing angle. Similar strengths and weaknesses are evident from the GPM observation of Typhoon Talim (Fig. 3.7), with our model particularly struggling to correctly classify the more discrete convective precipitation to the east and south of the eye (Fig. 3.8a,h). Our model also over-predicts convective precipitation in the north-northwest portion of the eyewall (Fig. 3.8g). Our model's over-prediction of convective precipitation here could be the result of randomness in the model training; or our model could be honed in on a particular predictor like the 89.0-GHz brightness temperature field, which shows a broad

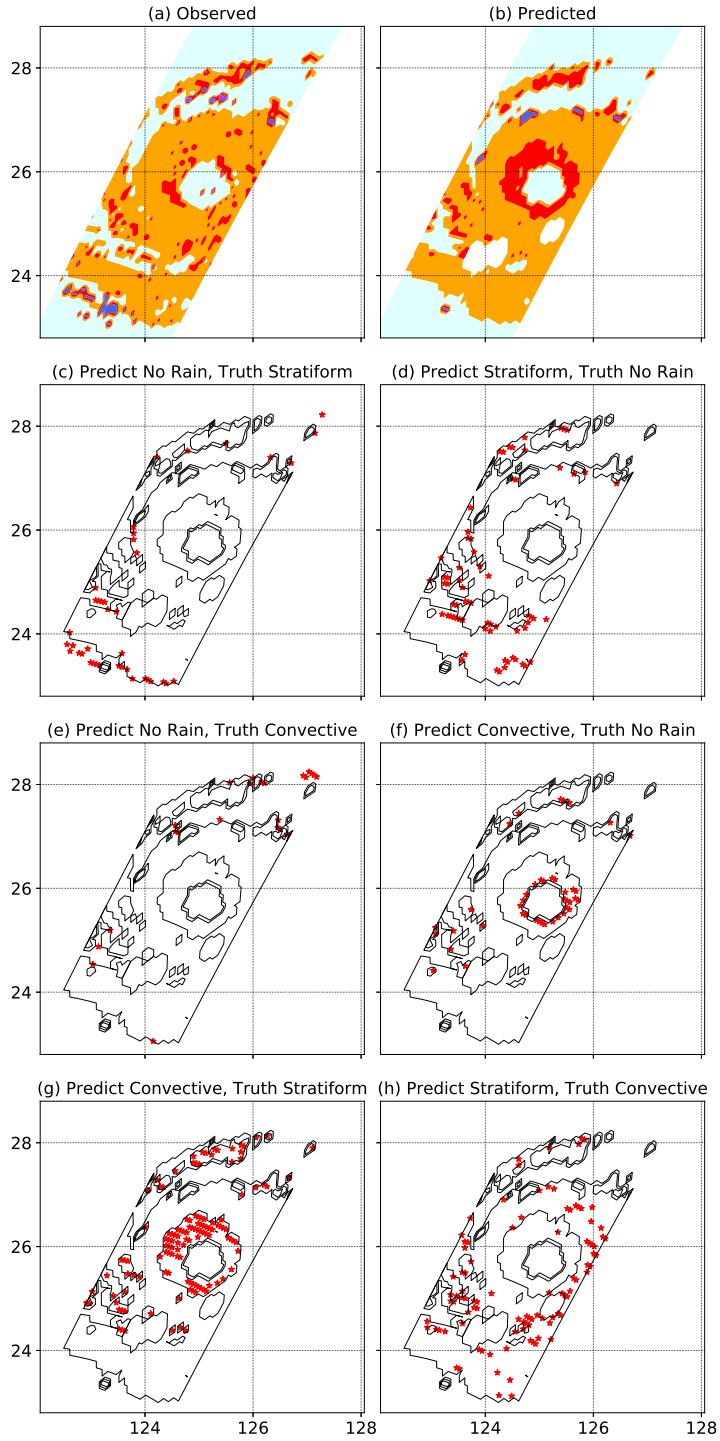

FIG. 3.8. Typhoon Talim's (a) DPR precipitation type classification re-gridded onto S1 as in Fig. 3.7e, (b) precipitation type classification from our random forest model as in Fig. 3.7f, and locations as marked by red stars of when our model (c) predicts no rain for a stratiform pixel, (d) predicts stratiform for a no rain pixel, (e) predicts no rain for a convective pixel, (f) predicts convective for a no rain pixel, (g) predicts convective for a stratiform pixel, and (h) predicts stratiform for a convective pixel.

band of cold brightness temperatures that is indicative of deep convection (Fig. 3.7b). And while we did not closely analyze the performance of the GPM DPR precipitation type classification, we also cannot rule out errors in the algorithm.

To investigate our model's overall performance on the 2017 observations, we calculate the F1 score and generate a confusion matrix (Table 3.4). The macro-averaged F1 score for the 2017 observations is 0.5555. Based on the range of baseline F1 scores that could result from internal model variability, which we discuss below, the F1 score of 0.5555 is significantly lower than 0.5730. This lower macro-averaged F1 score comes from over 2% decreases in the recall for stratiform precipitation, the precision for convecitve precipitation, and both recall and precision for both other and shallow precipitation types (Tables 3.3 and 3.4). The reduction in the macro-averaged F1 score for the 2017 observations points towards a slight problem with over-fitting of our model. The over-fitting is likely the result of our treatment of each pixel in our training and testing dataset as separate and independent. When we employed a randomization to split the training and testing dataset, and subsequently to split the training dataset into five folds for cross-validation and calibration, each split may contain pixels from the same set of observations (e.g. one passive microwave overpass of a tropical cyclone). Therefore, the train/test split and the cross-validation folds are not completely independent from one another, especially if there are overlapping pixels from the way we calculate the variability indices. Nonetheless, our calibrated model still performs well on the independent 2017 observations. Therefore, we will continue using the calibrated model in our discussions and offer further discussions on model over-fitting in the later sections.

To understand whether our model is overconfident or just *slightly more* confident whenever it makes a classification, we examine the distribution of model probability in various classification situations from the 2017 dataset. The sample size and frequency of each precipitation type for 2017 is similar to our testing dataset (Tables 3.2, 3.3, and 3.4). Like we have done so far, we will focus our discussions on no rain, stratiform, and convective precipitation types. For each correct and incorrect classification situation, we collect the model probability for no rain, stratiform and convective precipitation. We then group the model probabilities

into 0.05 probability bins and normalize their count by the sum of all bin counts so that their distribution represents relative frequencies.

Figure 3.9 shows the frequency of model probabilities in cases where our model is correct (diagonal white panels) and in cases where our model is incorrect (off-diagonal shaded panels). In cases where our model correctly classifies a precipitation type (Figs. 3.9a,e,i), the model probability for the correct precipitation type peaks on the higher end of the probability range (e.g., cyan line in (a), dashed orange line in (e), and dotted red line in (i)). In other words, when our model correctly classifies a precipitation type, it usually assigns a very high probability to that precipitation type. For example, 40% of the time when our model correctly classifies stratiform precipitation, it assigns an 80% probability of stratiform precipitation (Fig. 3.9e). Going from no rain to stratiform and convective precipitation, our model's confidence

TABLE 3.4. Confusion matrix for the calibrated model's predictions on the 2017 observations. The top number in each panel corresponds to the absolute number of cases that belong in that panel, with their row and column totals located in the far right column and bottom row, respectively. The top percentages in each panel is the recall and adds up to 100% in each row. The bottom and bold percentages in each panel is the precision and adds up to 100% in each column. The shaded cells highlight instances when our model correctly classifies a precipitation type.

		Predicted					
		No Rain	Stratiform	Convective	Other	Shallow	
Truth	No Rain	2,160,208 (97.9%) **(96.6%)**	25,210 (1.1%) **(9.3%)**	14,603 (0.7%) **(15.6%)**	4,187 (0.2%) **(32.2%)**	3,136 (0.1%) **(39.9%)**	2,207,344
	Stratiform	26,423 (9.6%) **(1.2%)**	217,838 (78.9%) **(80.5%)**	27,901 (10.1%) **(29.8%)**	3,559 (1.3%) **(27.4%)**	318 (0.1%) **(4.0%)**	276,039
	Convective	20,096 (21.7%) **(0.9%)**	22,156 (23.9%) **(8.2%)**	49,097 (52.9%) **(52.4%)**	89 (0.1%) **(0.7%)**	1,372 (1.5%) **(17.5%)**	92,810
	Other	14,312 (57.2%) **(0.6%)**	5,262 (21.0%) **(1.9%)**	296 (1.2%) **(0.3%)**	5,171 (20.7%) **(39.7%)**	0 (0%) **(0%)**	25,041
	Shallow	16,050 (76.5%) **(0.7%)**	45 (0.2%) **(0%)**	1,842 (8.8%) **(2.0%)**	3 (0%) **(0%)**	3,032 (14.5%) **(38.6%)**	20,972
		2,237,089	270,511	93,739	13,009	7,858	

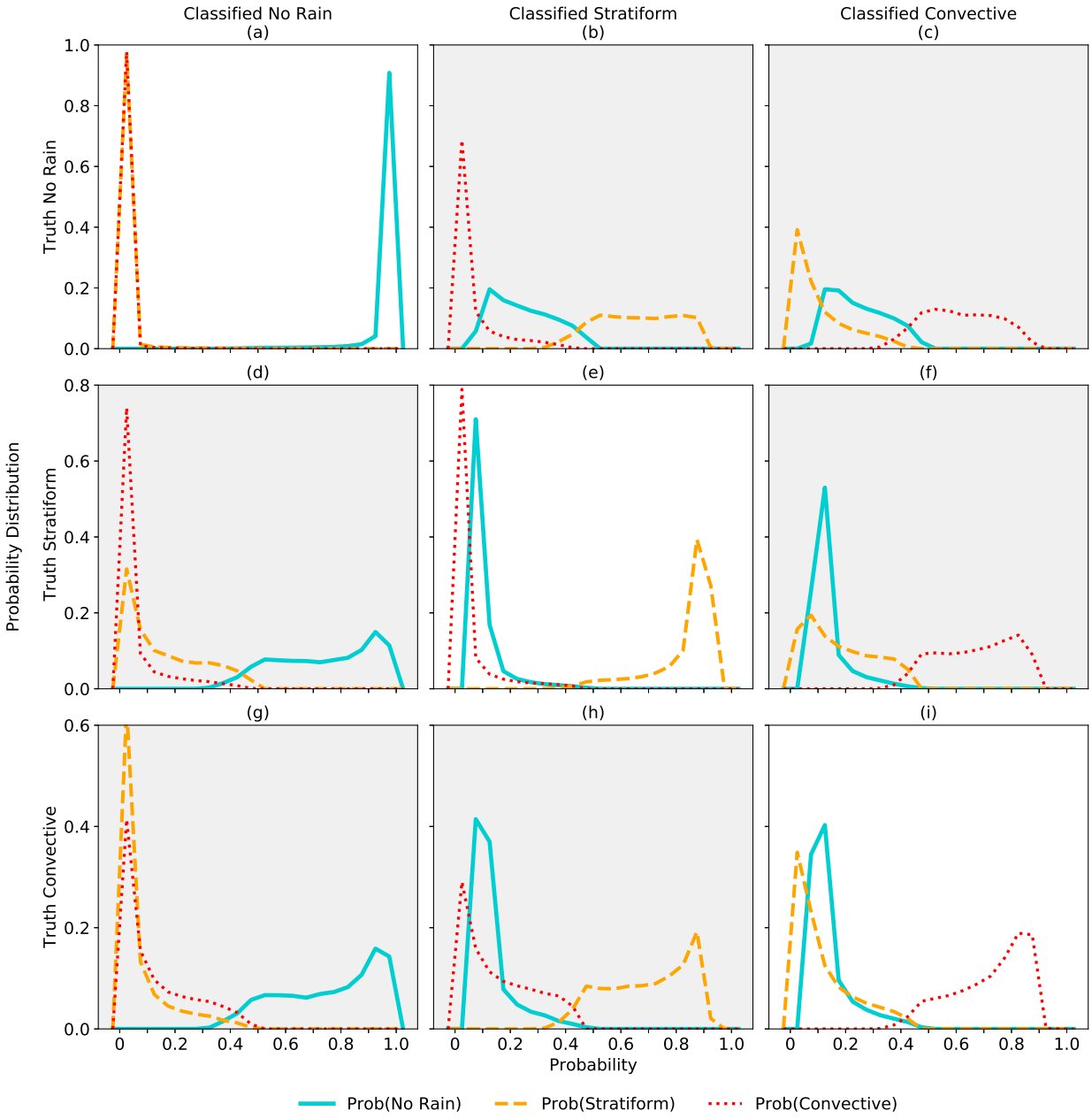

FIG. 3.9. Frequency distribution of model probabilities for GPM observations in 2017. The left, middle, and right column shows instances when our model classifies no rain, stratiform, and convection, respectively. The top, middle, and bottom row shows instances when the actual precipitation type is no rain, stratiform, and convective, respectively. The lines show the probability assigned to no rain (solid turquoise), stratiform (dashed orange), and convection (dotted red).

decreases, with an increasing frequency of lower probabilities (e.g. higher frequency of probabilities between 0.4 and 0.6 for correct classification of convective precipitation, compared to near-zero frequency of probability in the same range for correct classification of no rain).

When our model incorrectly classifies a precipitation type, the distribution of probabilities is wider for the classified precipitation type and the true precipitation type. For example, in cases when our model incorrectly classifies stratiform precipitation as no rain (Fig. 3.9d), the range of probabilities that our model assigned to no rain is wider compared to when our model correctly classifies no rain (Fig. 3.9a). The range of probabilities assigned to stratiform precipitation is also wider and extends to higher ranges of model probability — resulting in an overlap with the distribution of model probability for no rain. Meanwhile, the probability that our model assigns to convective precipitation is usually close to zero. Therefore, when our model incorrectly classifies stratiform precipitation as no rain, it is not as confident in its no rain classification. Our model also assigns the next highest probability to the correct precipitation type (stratiform), and is almost always confident that the precipitation type is not convective. This pattern is true for the incorrect classifications of all precipitation types shown here in Fig. 3.9.

We can also show this relationship in a different manner. Instead of showing the raw model probabilities, we repeat the steps above for probability *difference*. That is, in each of these correct and incorrect classification instances, we subtract the probability of the two other precipitation types from the probability of the classified precipitation type. A smaller probability difference indicate that our model is only marginally confident of its classification, while a larger probability difference indicate that our model is more confident of its classification. Figure 3.10 shows the distribution of probability difference for instances of correct and incorrect classifications.

Whenever our model correctly predicts no rain (Fig. 3.10a), the probability difference between no rain and stratiform or convective (e.g. probability of no rain minus the probability of either stratiform or convective) is most often above 0.95, which indicates that our model is justifiably confident in its classification of no rain. The same pattern is true for

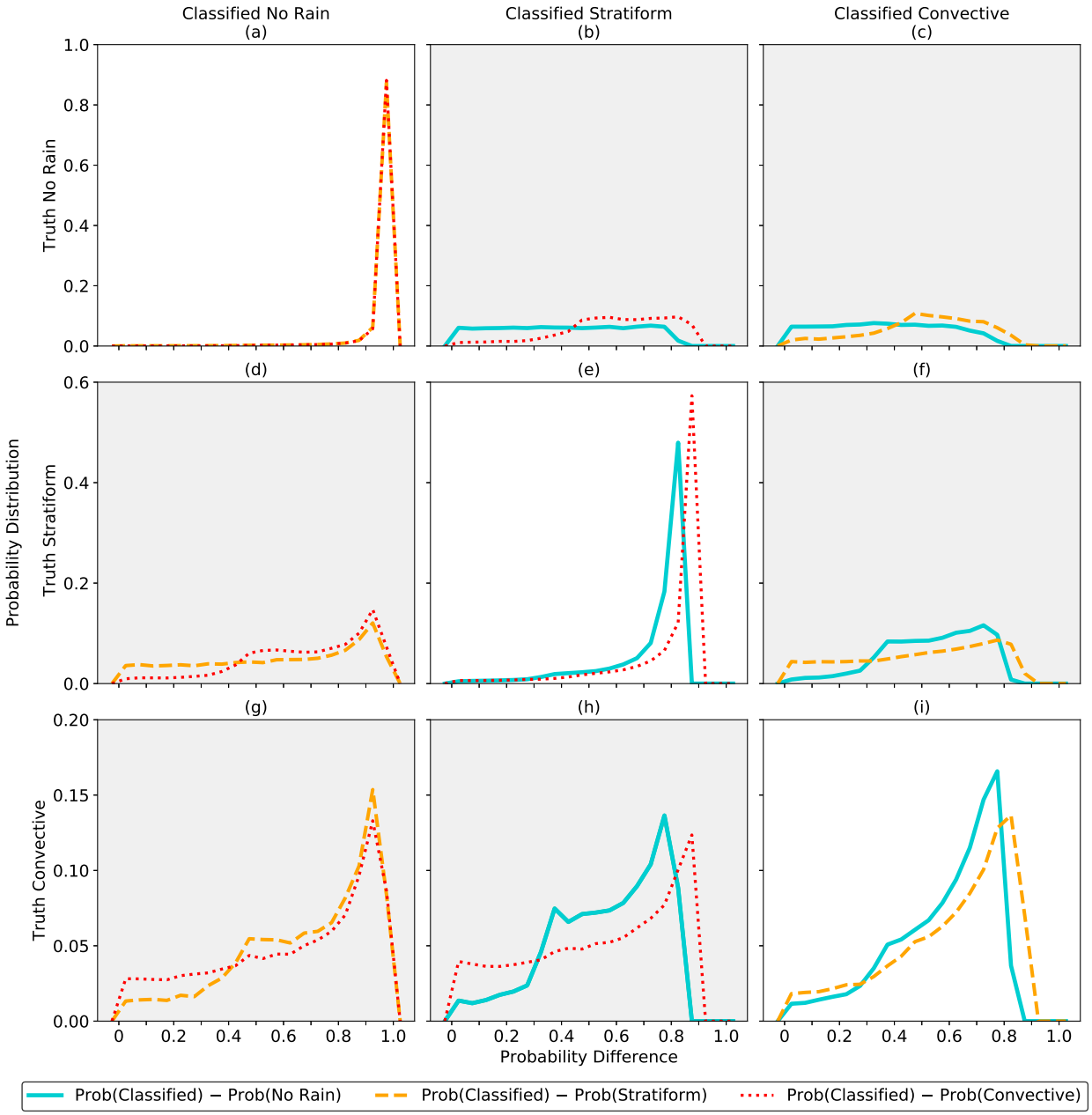

FIG. 3.10. Frequency distribution of probability difference between the classified precipitation type and the other two precipitation types for GPM observations in 2017. The left, middle, and right column shows instances when our model classifies no rain, stratiform, and convection, respectively. The top, middle, and bottom row shows instances when the actual precipitation type is no rain, stratiform, and convective, respectively. The lines show the probability difference between the classified precipitation type and: no rain (solid turquoise), stratiform (dashed orange), and convection (dotted red).

correct classifications of stratiform and convective precipitation (Figs. 3.10e,i). However, the increase in the frequency of smaller probability difference in correct classifications of stratiform and convective precipitation reflect the increasing uncertainty that we discussed above. When our model incorrectly classifies no rain as stratiform (Fig. 3.10b), the probability difference between stratiform and convective is most often above 0.5, which indicates that our model is usually confident that the precipitation type is not convective. More importantly, the probability difference between stratiform and no rain occurs over a wider range, with just as many large probability differences (e.g., >0.6) as do small probability differences (e.g., <0.2). The equal frequency of large and small probability differences indicate that in many cases where our model incorrectly classifies no rain as stratiform, it is only marginally confident that the precipitation type is stratiform, with the next likely class being the correct no rain precipitation type. The same pattern applies when our model incorrectly classifies no rain as convective (Fig. 3.10c).

When our model incorrectly classifies stratiform precipitation as no rain (Fig. 3.10d), the probability difference between no rain and convective is most often above 0.4, which means that our model is most often fairly and justifiably confident that the precipitation type is not convective. However, the most frequent probability difference between no rain and stratiform is 0.9, which indicates that our model is frequently overconfident in its incorrect classification of no rain for when the precipitation type is actually stratiform. Similar pattern applies to when our model incorrectly classifies convection as no rain (Fig. 3.10g), convection as stratiform (Fig. 3.9h), and, to a lesser extent, when our model incorrectly classifies stratiform precipitation as convective (Fig. 3.10f).

In summary, whenever our model correctly classifies no rain, it is extremely confident. When our model incorrectly classifies no rain as either stratiform or convective, it is less certain in its classification. While our model is less certain when it incorrectly classifies stratiform or convective precipitation — as marked by the more frequent occurrence of smaller probability difference between the classified and true classes — it also tends to be unjustifiably overconfident in its incorrect classification.

The probability analysis highlights two key points. First, non-raining scenes are often easier to classify than raining scenes. The more accurate classification of non-raining scenes is likely the result of non-raining scenes having markedly different passive microwave signatures than raining scenes. The model overconfidence in incorrect classification of stratiform and convective precipitation points towards more similar passive microwave signatures between the two precipitation types. But conventional wisdom suggest there should some degree of difference in the passive microwave signatures of stratiform and convective precipitation.

Therefore, the second key point from the probability analysis is that non-uniform beam filling is likely a factor in incorrect classification of raining scenes. Our classification approach does not account for multiple precipitation types within the GMI IFOV. While our approach to remap the GPM DPR precipitation type observation to S1 is based on representing the most dominant precipitation type, the most dominant passive microwave signature may be different. For example, our approach will classify a scene as stratiform if the scene contains a single convective pixel surrounded by stratiform precipitation. But the single convective pixel may be extremely deep convection, where the convective signatures in the passive microwave observations will dominate over the stratiform signatures. We briefly address our method of representing the precipitation type in the GMI IFOV in a later section.

3.4.2 Important Input Predictors

Now that we have some understanding of our model's performance in classifying precipitation type, we turn to understanding what input predictors are the most important for the best model performance. To evaluate the important input predictors, we apply a drop-column method. A drop-column method involves dropping a predictor, re-training, and evaluating the new model skill score relative to a baseline. The method is similar to the permutation importance in Scikit-Learn. Permutation importance involves permuting an input predictor, making a prediction, and evaluating the skill of the prediction relative to a baseline. However, the permutation importance routine in Scikit-Learn permutes only one input predictor at a time as it cycles through all of the input predictors. With our input predictors, two input predictors can co-vary, such as the vertical and horizontal polarization brightness temperatures

at a specific frequency. Permuting a single input predictor that is correlated with other predictors results in a biased assessment of the important input predictors. Therefore, we opt for the drop-column method. The drop-column method also allows us to customize the predictors that we want to drop in our model re-training, which we discuss below.

First, we need to obtain a baseline to understand whether a drop in model skill score is significant. To obtain a baseline, we re-train and calibrate our model with the same set of hyperparameters, but with different random seeds. The random seed determines which chunk of input data the model ingests to train each tree in the forest. By varying the random seed and analyzing the skill score, we can determine changes in the model skill score due to model internal variability. In the drop-column approach, any drop in model skill score outside of the range of model internal variability is thus significant. Due to limited computing resource, we re-train our model 15 times, with a random seed of 10 to 150 in increments of 10. Like in our model training and calibration, we use the F1 score as our scoring metric, with the lowest and highest F1 score being the range for a baseline score. We analyze the macro-averaged F1 score, as well as the F1 score for each precipitation type.

We first drop passive microwave brightness temperature observation from each GMI frequency. While the 36.64V and 36.64H channels are sensitive to emission by liquid precipitation over the radiometrically colder ocean surface, 36.64PCT is sensitive to scattering by large precipitation-sized ice particles. Therefore, we drop 36.64V and 36.64H separately from 36.64PCT. Conversely, 89.0V, 89.0H, and 89.0PCT are all sensitive to scattering by precipitation-sized ice particles. Therefore, we drop all of them simultaneously, to avoid leaving any correlated input predictors in our model re-training and re-calibration. We then individually drop the following derived input predictors: $EI_{36.64V}$, $VM_{36.64V}$, $VI_{89.0PCT}$, $VM_{89.0PCT}$, and $VC_{89.0PCT}$. Finally, we test our model's sensitivity to texture information in the ice-scattering channels by simultaneously dropping $VI_{89.0PCT}$, $VM_{89.0PCT}$, and $VC_{89.0PCT}$, followed by all of the texture information: $VM_{36.64V}$, $VI_{89.0PCT}$, $VM_{89.0PCT}$, and $VC_{89.0PCT}$.

Figure 3.11 shows the macro-averaged F1 score and the F1 score for each precipitation type from the drop-column approach. Note that re-training and re-calibrating our model with 15

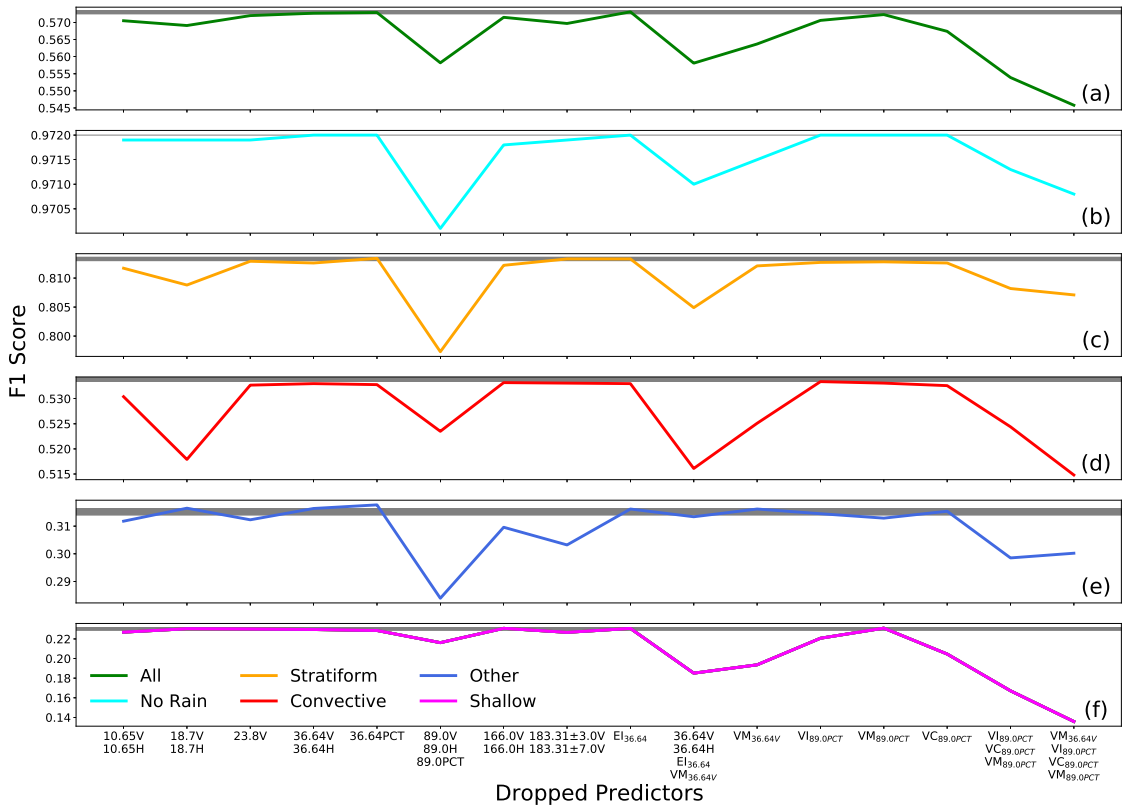

FIG. 3.11. Results from a drop-colum test showing the (a) macro-averaged F1 score, (b) F1 score for no rain, (c) F1 score for stratiform, (d) F1 score for convecitve, (e) F1 score for other, and (f) F1 score for shallow precipitation types. The gray bars in each panel shows the range of uncertainty due to model internal variability, with the model internal variability being a single value down to four decimal points (0.972) for no rain. We label the predictors we drop in the drop-column test at the bottom of the figure.

different random seeds may not be sufficient to obtain the reasonable range of baseline F1 scores. Therefore, while there are drops in F1 scores that are marginally outside our baseline, we focus our discussion on the more significant drops in F1 scores. While the drop-column approach tells us the most important predictors, it does not directly tell us how our model uses those predictors due to the non-linearity of the random forest model. As such, without further dissection, we can only hypothesize why a certain predictor is important. Nonetheless, using this drop-column approach, we can gain insight into the physicality of our model.

When we drop the 18.7-GHz observations, there is a slight decrease in the macro-averaged F1 score. This slight decrease is the result of a decrease in the F1 score for the respective stratiform and convective precipitation type (Fig. 3.11b,c). Our model's dependence on the 18.7-

GHz channel to detect stratiform and convective precipitation is unsurprising, since the 18.7-GHz channel is sensitive to emission by liquid water such as low-level clouds and rain. But the same dependence is not evident when we drop the 36.64-GHz observations. The 36.64-GHz observations are also sensitive to emission by liquid precipitation and has a finer IFOV than the 18.7-GHz channel. Therefore, the 36.64-GHz observations should have similar importance to our model's predictions as the 18.7-GHz channel, if not more.

To investigate the apparent lack of importance of the 36.64-GHz channel, we drop all of the emission-based predictors from the 36.64-GHz channel. They are: 36.64V, 36.64H, $EI_{36.64}$, and $VM_{36.64V}$. The macro-averaged F1 score drops significantly when we drop all of the emission-based predictors from the 36.64-GHz channel, which indicates that the 36.64V and 36.64H co-varies with $EI_{36.64}$ such that the presence of one is sufficient to make up for the absence of the other. In other words, $EI_{36.64}$ is a redundant predictor. Note however, that the drop in the macro-averaged F1 score when we drop all of the emission-based predictors from the 36.64-GHz channel is just slightly more than the macro-averaged F1 score when we drop the 36.64-GHz texture information alone ($VM_{36.64V}$). Therefore, while 36.64V, 36.64H, and $EI_{36.64}$ are important, $VM_{36.64V}$ alone contains more useful information for our model.

The macro-averaged F1 score drops when we drop the 89.0-GHz observations. The drop in F1 score is also evident across all precipitation types, but to a much lesser extent in the shallow rain prediction. The 89.0-GHz is sensitive to scattering of terrestrial radiation by large precipitation-sized ice particles that typically occurs in deep convection. However, the F1 scores for the other precipitation types also drop. We hypothesize that the drop in the F1 score for the non-convective precipitation type is the result of our model's inability to distinguish convective precipitation from the remaining precipitation type.

There is a slight decrease in the macro-averaged F1 score when we drop the 183.31-GHz observations. The drop in the macro-averaged F1 score comes from a drop in the F1 score for the other precipitation type. The 183.31-GHz is sensitive to attenuation by atmospheric water vapor such that terrestrial emissions in this channel comes mainly from near the top of the water vapor layer. The presence of liquid or ice cloud can also attenuate terrestrial

radiation at this frequency. If we are correct in surmising that the other precipitation type is essentially overhanging precipitation or cirrus, then our model's dependence on the 183.31-GHz channel is reasonable. The emission-based channels would not be able to detect the other precipitation type because it would not have a low-level emission signature. Similarly, if the other precipitation type consists of only cloud ice and not *precipitation-sized* ice particles, the 89.0-GHz would not be able to detect it either.

When we individually drop the ice-scattering texture information ($VI_{89.0PCT}$, $VM_{89.0PCT}$, and $VC_{89.0PCT}$), the F1 score across the different precipitation types drop only marginally, except the shallow rain type, which dropped noticeably when we drop $VC_{89.0PCT}$. When we drop all of the ice-scattering texture information, the F1 scores drop significantly. Similar to the 36.64-GHz emission predictors, the presence of any two combination of ice-scattering texture information is sufficient to make up for the absence of one. The biggest drop in the macro-averaged F1 score occurs when we drop all of the texture information. As we discussed earlier, the passive microwave texture provides information on the smoothness of the cloud and precipitation fields. While Anagnostou and Kummerow (1997) and Hong et al. (1999) developed these texture indices to delineate convective from stratiform precipitation, our model may also use it to delineate convective precipitation from no rain, or to detect stratiform precipitation from the smoothness of the texture information. Nonetheless, our drop-column test highlights texture information is the most important predictor in classifying precipitation type from passive microwave observations.

To reiterate, we cannot know with certainty how our model treats each predictor in making its predictions unless we delve deeper and print out the logic of each tree. Nonetheless, since our model placed its importance on the emission channels (18.7- and 36.64-GHz), the ice-scattering channels (89.0-GHz), and their corresponding texture information, we know that our model is making its predictions based on the physical principles that those predictors possess.

3.4.3 Coarser Resolution Data

Recall earlier that our model performance is likely dependent on how we choose to represent the precipitation type in the GMI IFOV. Here, we test the model performance for a coarser representation of precipitation. For our initial model, we re-grid the DPR precipitation type onto S1 with FOV = 5 km for the weighting function in an attempt to retain extra detail in the precipitation field while representing the most dominant precipitation type in the GMI IFOV. To test our model's sensitivity to a coarser resolution data, we repeat the GridSearchCV, training, and calibrating steps using FOV = 25 km in the weighting function when re-gridding the DPR precipitation type onto S1. Using FOV = 25 km will weight the dominant precipitation type over a larger area, essentially smoothing the precipitation field. We keep everything else the same — using 50 trees and a balanced class weight, separating our input data into 5 folds for cross-validation, and using the macro-averaged F1 score as our scoring metric in the GridSearchCV step. The best model hyperparameters are a tree depth of 70 and a minimum number of samples in a leaf node of 5. Since a tree depth of 70 is at the end of our initial search range, we extend the search range to include tree depth of 80 and 90, but find that the best tree depth is still set at 70. Using 50 trees, a depth of 70, and a minimum number of samples in a lead node of 5, we re-train and re-calibrate our model. Table 3.5 shows the confusion matrix for the calibrated model with a smoother precipitation field.

Except for the precision of the shallow precipitation type, the recall and precision of each precipitation type increase relative to our earlier model (Tables 3.5,3.3). The general increase in recall and precision for each precipitation type indicates that the coarser precipitation field is more representative of the input predictors, since some important input predictors have a coarse IFOV (e.g. an ellipse of approximately 10.9 × 18.1 km for the 18.7-GHz channel). However, the increase in recall and precision is relatively minor, which indicates the importance of accounting for non-uniform beam filling to improve model performance. In other words, while selecting the most dominant precipitation type over a larger area leads to a marginally better model performance, a different representation of precipitation — one that accounts for the mixture of precipitation types within the GMI IFOV — will more likely lead to a more

TABLE 3.5. Confusion matrix for the calibrated model with a smoother precipitation type field. The top number in each panel corresponds to the absolute number of cases that belong in that panel, with their row and column totals located in the far right column and bottom row, respectively. The top percentages in each panel is the recall and adds up to 100% in each row. The bottom and bold percentages in each panel is the precision and adds up to 100% in each column. The shaded cells highlight instances when our model correctly classifies a precipitation type.

		Predicted					
		No Rain	Stratiform	Convective	Other	Shallow	
Truth	No Rain	2,361,684 (98.1%) **(97.7%)**	27,204 (1.1%) **(8.1%)**	13,612 (0.6%) **(14.6%)**	3,607 (0.1%) **(27.8%)**	1,716 (0.1%) **(37.5%)**	2,407,823
	Stratiform	23,957 (7.2%) **(1.0%)**	281,932 (84.8%) **(84.0%)**	23,385 (7.0%) **(25.1%)**	3,089 (0.9%) **(23.8%)**	93 (0%) **(2.0%)**	332,456
	Convective	14,509 (15.7%) **(0.6%)**	22,090 (23.9%) **(6.6%)**	54,932 (59.3%) **(58.9%)**	39 (0%) **(0.3%)**	1,021 (1.1%) **(22.3%)**	92,591
	Other	10,495 (49.6%) **(0.4%)**	4,279 (20.2%) **(1.3%)**	145 (0.7%) **(0.2%)**	6,255 (29.5%) **(48.2%)**	0 (0%) **(0%)**	21,174
	Shallow	6,124 (67.3%) **(0.3%)**	15 (0.2%) **(0%)**	1,204 (13.2%) **(1.3%)**	0 (0%) **(0%)**	1,750 (19.2%) **(38.2%)**	9,093
		2,416,769	335,520	93,278	12,990	4,580	

significant improvement in model performance. Instead of selecting the most dominant precipitation type within the GMI IFOV like we have done here, future precipitation type classification may instead choose to have a separate class of "mixed" precipitation types, or train a model to regress the percent coverage of the different precipitation types in the GMI IFOV (e.g., 5% no rain, 45% stratiform, and 50% convective).

3.4.4 Quantifying Model Probabilities

In analyzing the probability difference in our model, we showed that our model has a range of probability differences between the predicted class with the actual class in both correct and incorrect predictions. Therefore, quantifying the probability of a correct detection based on our model's own probability may be misleading. Instead, we propose using the precision calculation as a general model probability. The precision is the fraction of correct classifications relative all of the classifications a model makes. Extending this definition to include other

precipitation types, we can quantify the probability in our model's predictions. For example, whenever our model predicts convective precipitation, there is a 54.7% chance that the pixel is actually convective and an approximately 27.5% chance that the pixel is actually stratiform (Table 3.3). Having this value provides some quantitative information for us to use this precipitation type classification model to investigate the distribution of stratiform and convective precipitation in tropical cyclones.

3.5 Conclusion

Using a random forest classifier, we have developed a precipitation type classification algorithm for tropical cyclones over the ocean that uses GMI brightness temperature observations as input. Our model correctly classifies 97.7% of no rain pixels, 81.9% of stratiform pixels, and 52.2% of convective pixels. Whenever our model classifies no rain, stratiform, and convective pixels, there is a respective 96.7%, 80.7%, and 54.7% chance that it correctly classifies the precipitation type. While our model has a more difficult time separating convective from stratiform precipitation, it does extremely well in delineating significantly precipitating regions of stratiform and convective precipitation from no rain.

Our model exhibits difficulty in correctly classifying precipitation type on the edges between two or more precipitation types. In addition, our model is relatively overconfident when it misclassifies stratiform and convective precipitation, but just slightly more confident when it misclassifies no rain. Our drop column-test shows that our model leverages observations from the liquid emission channels (18.7-GHz and 36.64-GHz) and ice-scattering channel (89.0-GHz) to make its predictions. However, the most important input predictor for the best precipitation type classification is the texture information. Finally, we re-train and re-calibrate our model for a smoother precipitation type field and found increases in model scores. But the increases in model scores are marginal, indicating the need for a different method to represent the precipitation types in the GMI IFOV. Future approaches to precipitation type classification from passive microwave observations may include an additional "mixed" precipitation type within the passive microwave IFOV, or the regression of the percent coverage of multiple precipitation types within the passive microwave IFOV. Such approaches must also be mindful

to not over-smooth the precipitation type field to the point where important features like the eye of a small tropical cyclone cannot be distinguished.

Our model is slightly over-fitted as our train/test dataset and cross-validation folds do not contain completely independent observations. One of the ways we can reduce this over-fitting is to manually separate the data ourselves such that the train/test split and the cross-validation folds are completely independent. For example, we can use the GPM observations from the 2014–2018 tropical cyclone seasons to make up each of the five cross-validation folds and test our model on the GPM observations from the 2019 tropical cyclone season. We can also leverage other random forest hyperparamters, such as the minimum number of samples at a split. Nonetheless, even with a slight overfitting, our model did relative well on an independent dataset.

While we trained our model on GMI observations of tropical cyclones, our method is applicable to precipitation type classification of any precipitating systems over the ocean. Generating a precipitation type classification for precipitating systems over land may require a more rigorous approach to consider the different land surface emissivities, but not impossible (Petković et al. 2019). The significance of having texture information to predict precipitation type highlights the need to use methods that can directly ingest the texture information, such as a convolutional neural network, which does not require a calculation to represent texture as we have done in our approach.

CHAPTER 4

CONVECTION AND PRECIPITATION DURING TROPICAL CYCLONE INTENSITY CHANGE

4.1 INTRODUCTION

The scientific community has made significant advances toward understanding the factors that impact tropical cyclone intensity change. For example, tropical cyclone intensification requires a favorable environment that includes sufficiently warm sea surface temperatures, low vertical shear of the horizontal winds (hereafter, vertical shear), and high tropospheric humidity. However, within these favorable environments, tropical cyclone intensification rates can still vary significantly (Hendricks et al. 2010). The varying intensification rates for tropical cyclones in favorable environments indicate the importance of tropical cyclone internal dynamics. Among the various tropical cyclone internal dynamical mechanisms, the exact role of precipitation types (i.e., shallow, moderate, or deep convective precipitation, or stratiform precipitation) on intensification is still unclear and requires further research.

The distribution of precipitation in tropical cyclones generally depends on the environmental vertical shear (Hence and Houze 2011; Reasor et al. 2013; DeHart et al. 2014). In the tropical cyclone inner core, convection typically initiates in the downshear right quadrant. As convection travels cyclonically downwind and continues to mature, precipitation reach maximum in the downshear left quadrant. In the upshear left quadrant, precipitation intensity decreases as sinking motion dominates, with the occurrence of convection and precipitation reaching a minimum in the upshear right quadrant. This shear-induced asymmetry affects how convection and precipitation may interact with the tropical cyclone inner core, and subsequently influence intensity change.

Some studies associate tropical cyclone intensification with the occurrence of deep convection. For example, using a multi-platform observational analysis, Guimond et al. (2010) highlighted a period of episodic convective bursts that preceded the rapid intensification of Hurricane Dennis (2005). Using composite airborne Doppler radar observations, Rogers et al. (2013) showed that intensifying tropical cyclones have a higher occurrence of convective bursts

relative to steady-state tropical cyclones. The convective bursts occur most frequently within the radius of maximum winds in intensifying tropical cyclones, and outside the radius of maximum winds in steady-state tropical cyclones. Deep convection also preceded the rapid intensification of Hurricane Earl (2010) and Edouard (2014) (Stevenson et al. 2014; Rogers et al. 2015; Susca-Lopata et al. 2015; Rogers et al. 2016).

These studies hypothesize that convection contributes towards tropical cyclone intensification through any combination of a few mechanisms (Shapiro and Willoughby 1982; Schubert and Hack 1982; Vigh and Schubert 2009; Smith et al. 2009; Stevenson et al. 2014; Rogers et al. 2020). First, the strong subsidence on the inner edge of deep convection effectively warms and dries the tropical cyclone warm core, which — through thermal wind balance — would lead to intensification. Second, the occurrence of deep convection is crucially within the high inertial stability region of the tropical cyclone inner core. The high inertial stability produces stronger resistance to radial parcel displacements, leading to a configuration that is more efficient for convective heating to contribute towards the spin-up of the tangential winds. Third, deep convection helps align the lower- and upper-level vortex, which helps consolidate the tropical cyclone for intensification. Finally, deep convection in the tropical cyclone inner core effectively converges angular momentum within the boundary layer faster than it dissipates through friction, thus increasing the tangential wind speed.

Note that for these mechanisms to contribute towards intensification, convection must occur near the radius of maximum winds. In addition, these mechanisms are also more effective when the initially asymmetric convection and its attendant kinematic and thermodynamic fields axisymmetrize around the inner core, resulting in a more axisymmetric influence on the tropical cyclone. The more axisymmetric distribution of convection in intensifying tropical cyclones manifests themselves in shear-relative quadrants. Wadler et al. (2018) showed that, compared to steady-state tropical cyclones, intensifying tropical cyclones have significantly more convective bursts that are stronger and reach a higher altitude in the upshear left quadrant. They hypothesize that the presence of convective bursts in the upshear left quadrant of intensifying tropical cyclones represents the ability of downshear convection

to persist as it rotates cyclonically downwind into the upshear quadrants, resulting in a better projection of convective heating into the azimuthal mean.

Other studies showed that the occurrence of deep convection is not necessarily a precursor for tropical cyclone intensification. Using convective proxies from the Tropical Rainfall Measurement Mission (TRMM) Precipitation Radar (PR) observations of the most intense convection in the inner core of tropical cyclones, Jiang (2012) showed that the intensity of convection increases with increasing future intensification rate. However, when she used the convective proxies to predict rapid intensification, the increase in prediction skill was relatively marginal. Meanwhile, Kieper and Jiang (2012) analyzed the 37-GHz passive microwave color composite satellite imagery from the Naval Research Laboratory and found that a ring of shallow liquid precipitation, which occurs in the form of a cyan ring in the imagery, usually occurs around the time of the strongest 24-hour intensification rate in rapid intensification events.

Tao and Jiang (2015) looked at the distribution of shallow to very deep precipitation in tropical cyclones and found that the onset of rapid intensification coincides with a more widespread and symmetric shallow to moderately deep precipitation in the tropical cyclone inner core. They hypothesized that the contribution of latent heating from these more symmetric shallow to moderately deep precipitation is more important for the initiation and maintenance of rapid intensification compared to the contribution from asymmetric deeper convection, similar to the findings of Zagrodnik and Jiang (2014). Tao et al. (2017) further found that stratiform precipitation increases in the hours prior to the onset of rapid intensification, particularly in the upshear quadrants. They hypothesized that the increased stratiform coverage moistens the upper levels, leading to a more symmetric rainfall. Similarly, using passive microwave brightness temperature thresholds as proxies for precipitation and deep convection, Alvey et al. (2015) found that while tropical cyclones with stronger intensification rates contain more deep convection, the increase in occurrence and symmetry of moderate precipitation correlates better with future intensity change. In addition, they showed that the occurrence of deep convection did not noticeably increase prior to the onset of intensification, and instead increases after the onset.

In summary, the authors in these latter set of studies hypothesize that i) an increase in symmetry of shallow to moderately deep precipitation correlates better with future intensity change, particularly rapid intensification, ii) the latent heating from the shallower and more symmetric inner core precipitation contributes more significantly to intensification than do asymmetric deeper convection, iii) the more symmetric shallow to moderately deep precipitation also pre-conditions the environment for deeper convection, and iv) the presence of widespread deep convection in the inner core of rapidly intensifying tropical cyclones is a symptom of rapid intensification — not a cause. They did note that when deep convection does occur prior to or at the onset of intensification, they tend to occur closer to the inner core and in the upshear left quadrant compared to the other intensity change categories (i.e. steady-state or weakening). These hypotheses are not necessarily in conflict with the findings of Rogers et al. (2013, 2016), who showed an increase in precipitation symmetry alongside an increase in deep convection in intensifying tropical cyclones.

Tao and Jiang (2015) offered an explanation to reconcile the two conflicting hypotheses surrounding the role of deep convection in tropical cyclone intensity change. They suggested that the conflicting hypotheses result from the definition of the onset of intensification. For example, the observations of Hurricane Dennis in Guimond et al. (2010) occurred when rapid intensification was already happening, while the observations of Hurricane Earl (Stevenson et al. 2014; Rogers et al. 2015; Susca-Lopata et al. 2015) occurred following a period of slow intensification, when isolated deep convection can occur.

A final set of studies instead proposed a more synergistic role of deep convection and shallower precipitation in intensification. In a simulation of the rapid intensification of Hurricane Dennis (2005), Rogers (2010) showed that the first indication of rapid intensification occurred in the form of increased vertical mass flux in the low levels, which was accomplished mainly by weak low-level updrafts and the background tropical cyclone secondary circulation. Deep convective bursts contribute towards this increase in vertical mass flux, which enhances the tropical cyclone secondary circulation, and subsequently enhances the tropical cyclone primary circulation. The tropical cyclone reaches a tipping point where the increasing inner-core

inertial stability from the intensifying primary circulation is sufficiently high enough such that the occurrence of the next convective burst within the inner core results in the most efficient conversion of latent heating to kinetic energy, and rapid intensification occurs.

The synergy between deep convection and shallower convection can also occur in a different manner. In the intensification of Hurricane Hermine (2016) from a tropical depression to a category 1 hurricane, Rogers et al. (2020) showed that deep convection helped align the initially displaced low-level and mid-level circulation centers through vorticity stretching and tilting. After Hermine achieved vertical alignment, the vertical mass flux profile transitioned into a bottom heavy profile and coincides with the presence of widespread shallower convection. The widespread shallower convection stretched and tilted vorticity over a larger scale, which helped sustain the alignment over a larger spatial and longer temporal scale.

To investigate the distribution of convection and precipitation in tropical cyclone intensity change, we use a precipitation type classification designed for the Global Precipitation Measurement (GPM) satellite. We obtain GPM observations of tropical cyclones from the Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TCPRIMED; Razin et al. 2021). In Chapter 3, we generated the precipitation type classifier using a random forest model and the GPM passive microwave brightness temperature observations as input predictors. Since this precipitation type classifier uses the GPM passive microwave observations, it can observe a larger section of the tropical cyclone compared to the narrower swath of the GPM dual-frequency precipitation radar, providing a different view of the convection and precipitation distribution in tropical cyclones. In section 2, we present our methods in selecting and compositing GPM observations. We then present out results in section 3 and discuss its significance in section 4. In section 5, we offer concluding remarks.

4.2 METHODS

4.2.1 TC PRIMED

For this study, we use GPM observations of tropical cyclones from the Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED; Razin et al. 2021).

TC PRIMED is a dataset of satellite passive microwave observations centered on tropical cyclones and contains other data such as retrieved precipitation from NASA's Goddard Profiling algorithm (GPROF), coincident infrared brightness temperatures, TRMM and GPM precipitation radar observations, tropical cyclone information such as best-track location and intensity, and derived environmental diagnostics and three-dimensional fields from the fifth version of the European Center for Medium-Range Weather Forecasts (ECMWF) re-analysis (ERA5). TC PRIMED uses a consistent definition of intensity across all tropical cyclone basins, which is the maximum 1-minute sustained tangential wind speed at 10-meter altitude. While TC PRIMED contains tropical cyclone observations from various satellite passive microwave sensors, we use specifically the GPM observations since we have developed a precipitation type classifier for the GPM sensor, which we summarize below. In addition, we use the environmental diagnostics and best-track information from TC PRIMED to isolate tropical cyclones in favorable environments and separate GPM observations into the different intensity change groups.

4.2.2 Precipitation Type Classification

In Chapter 3, we trained a random forest model to classify precipitation type using GPM passive microwave observations. The model classifies five precipitation types — i) no rain, ii) stratiform, iii) convective, iv) other/cirrus, and v) shallow rain — using input predictors that include brightness temperature observations from all of the GPM channels, derived brightness temperature quantities like the 89.0-GHz polarization-corrected brightness temperatures (PCT), and texture information from the 89.0-GHz and 36.64-GHz channels. The model performs exceptionally well in delineating no rain from significant precipitation like stratiform and convective precipitation, providing a great representation of precipitation coverage in tropical cyclones. This precipitation type classification approach contrasts with the approach of Alvey et al. (2015), who used a brightness temperature threshold as proxy for precipitation.

The model struggles to identify the other and shallow precipitation types. Since the other and shallow precipitation types are typically small, infrequent, and have no significant impacts on tropical cyclones, we will not discuss their distributions. As we will show later, the

dominant forms of precipitation in our composites are convective and stratiform precipitation. Overall, the model correctly detects over 96%, 80%, and 50% of non-raining, stratiform, and convective precipitation, respectively. We showed in Chapter 3 that much of the model failures come from the detection of isolated or randomly distributed convection. Conversely, the model can detect the more organized convection of the eyewall and rainbands relatively well.

4.2.3 Selection of Tropical Cyclone Observations

For each GPM tropical cyclone overpass in TC PRIMED, we use the closest synoptic time (00,06,12,18 UTC) to characterize the tropical cyclone's intensity and environmental conditions from the best-track and diagnostics data. Then, we group tropical cyclone observations into six different intensity change *groups*. They are weakening (WK), steady-state (SS), and intensifying (IN) tropical cyclones for the minor and major intensity *category*. Minor tropical cyclones include category 1 and 2 tropical cyclones on the Saffir–Simpson Hurricane Wind Scale, while major tropical cyclones include category 3 and above. We do not include tropical depressions or tropical storms to focus our analysis on the distribution of convection and precipitation in more developed tropical cyclones. Since the best-track intensity information has a resolution of 5 knots, we define intensity change based on their future 12-hour intensity change where WK tropical cyclones weaken by 10 knots or more ($dV(+12) \leq -10$ kts), SS tropical cyclones remain between -5 and 5 knots ($-5 \leq dV(+12) \leq 5$ kts), and IN tropical cyclones strengthen by 10 knots or more ($dV(+12) \geq +10$ kts). The tropical cyclones that we sample must also remain at least 150 km away from land throughout this 12-hour period. In addition, to remove the effects of eyewall replacement cycles, we manually inspect and remove any GPM observations of tropical cyclones with concentric eyewalls. However, this approach does not account for observations of tropical cyclones with a single eyewall prior to or after eyewall replacement cycles.

To isolate the impact of convection and precipitation distribution on intensity, we composite the GPM observations of tropical cyclones only in favorable environments. We define favorable environments as sea-surface temperatures equivalent to or greater than 27°C and

environmental vertical shear of equivalent to or less than 10 m s^{-1}. The criteria include environments that are marginally favorable, since a more strict criteria representing the most favorable environments will reduce our sample size. We do not further isolate the tropical cyclones based on the environmental relative humidity, but a brief analysis of the averaged 850–500 mb relative humidity shows that they are similar across the intensity change groups. Table 4.1 shows the mean and standard deviation of the tropical cyclone and environmental characteristics sampled in each intensity change group, which also includes the intensity change in the 12 hours prior to the observation time (dV(−12)).

4.2.4 Compositing

For each GPM tropical cyclone overpass, we rotate the satellite observation such that the environmental vertical shear vector points upward in our analysis (along the +y axis). We mirror the satellite observations of Southern Hemisphere tropical cyclones about the equator to ensure consistent representation of the shear-relative convection and precipitation distribution between the Northern and Southern Hemisphere tropical cyclones. Finally, to account for different tropical cyclone sizes, we apply a radius scaling using the scaling parameter from Knaff et al. (2014, 2017). Using this scaling parameter, a tropical cyclone with a climatologically small wind field will have a scaled radius that is larger than its physical radius. Conversely, a tropical cyclone with a climatologically large wind field will have a scaled radius that is smaller than its physical radius. In theory, this scaling parameter maps the satellite observations from large and small tropical cyclones into radial regions with similar surface wind speeds. Subsequently, we invert the shear-relative azimuth and scaled radius of the observations into scaled zonal and meridional distances from the tropical cyclone center (x* and y*, respectively.)

We then group the satellite observations into the scaled 20 km* × 20 km* bins. The satellite observations include the precipitation type, mean surface precipitation rates, and 89.0-GHz PCT. Figures 4.1 and 4.2 show the number of tropical cyclones and satellite pixels sampled in each bin. Note that the number of tropical cyclones sampled in Fig. 4.1 does not quite match that in Table 4.1. In addition, the number of tropical cyclones and pixels sampled

TABLE 4.1. The mean and standard deviation (in parentheses) of the tropical cyclone and environmental characteristics in each intensity change group.

	Minor			Major		
	WK	SS	IN	WK	SS	IN
Number of Tropical Cyclones	36	80	128	65	61	44
Intensity / knots	81.39 (10.86)	82.50 (10.03)	77.38 (10.39)	121.77 (15.37)	119.02 (13.06)	115.68 (12.88)
dV(−12) / knots	−3.89 (9.19)	5.19 (13.32)	14.8 (9.10)	3.23 (15.29)	8.93 (16.59)	20.11 (11.59)
dV(+12) / knots	−14.44 (7.44)	0.88 (4.04)	21.33 (11.63)	−14.46 (6.38)	−0.16 (4.18)	17.05 (9.23)
SST / °C	28.05 (0.78)	28.43 (0.82)	28.47 (0.86)	28.23 (0.80)	28.61 (0.76)	28.66 (0.65)
Vertical Shear / m s^{-1}	6.1 (2.47)	5.28 (2.24)	4.82 (2.47)	5.24 (2.53)	4.83 (2.28)	4.18 (2.0)
850−500 mb RH / %	73.29 (8.03)	74.26 (7.82)	76.41 (8.53)	75.08 (7.78)	76.51 (7.44)	77.03 (6.42)

has a maximum near the tropical cyclone center. This slight discrepancy results from some non-overlapping GPM overpasses. In this study, minor IN tropical cyclones are the most sampled intensity change group, followed by minor SS tropical cyclones (Table 4.1). The major IN group has a slightly larger sample size than the minor WK group, with the major SS and major WK groups having a similar sample size. The least sampled intensity change group is the minor WK tropical cyclones. Therefore, we must be cautious when interpreting composites for the minor WK tropical cyclones.

For each intensity change group, we average the surface precipitation rate in each 20 km* × 20 km* bin. For the precipitation type variable, we calculate the fractional occurrence of rainfall, which includes stratiform, convective, other, and shallow precipitation types, in the same bin. Subsequently, we calculate the fractional occurrence of each precipitation type for precipitating regions. For the 89.0-GHz PCT, we obtain the 5th-percentile values as proxy for the intensity of the strongest convection in each bin. To highlight the differences between the intensity change groups, we separately average the rainfall occurrence, convective fraction, and surface precipitation rates of the WK, SS, and IN groups in their respective intensity category (i.e. minor and major). The relative differences of the variables that we show in this

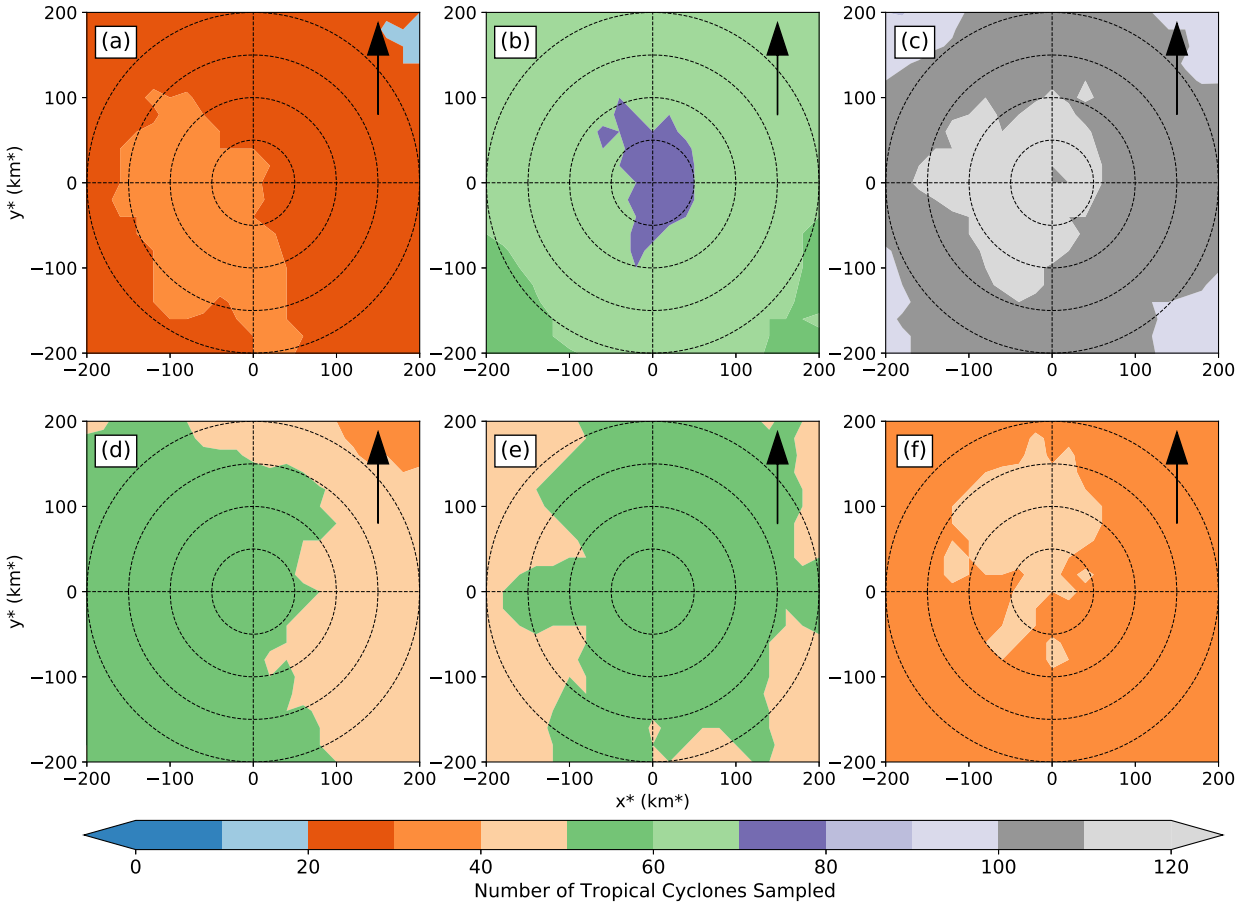

FIG. 4.1. Number of tropical cyclones sampled in each scaled 20 km* × 20 km* bins for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

paper will therefore be with respect to the mean of their respective intensity category. By averaging the variables after we have calculated them for the individual WK, SS, and IN group, we remove the bias from the uneven sample sizes between the groups.

4.3 RESULTS

4.3.1 Tropical Cyclone and Environmental Characteristics

From Table 4.1, minor WK and SS tropical cyclones have similar mean intensities, while minor IN tropical cyclones have a lower mean intensity at the time of the GPM observation. Minor WK tropical cyclones were, on average, already weakening in the 12 hours prior to the

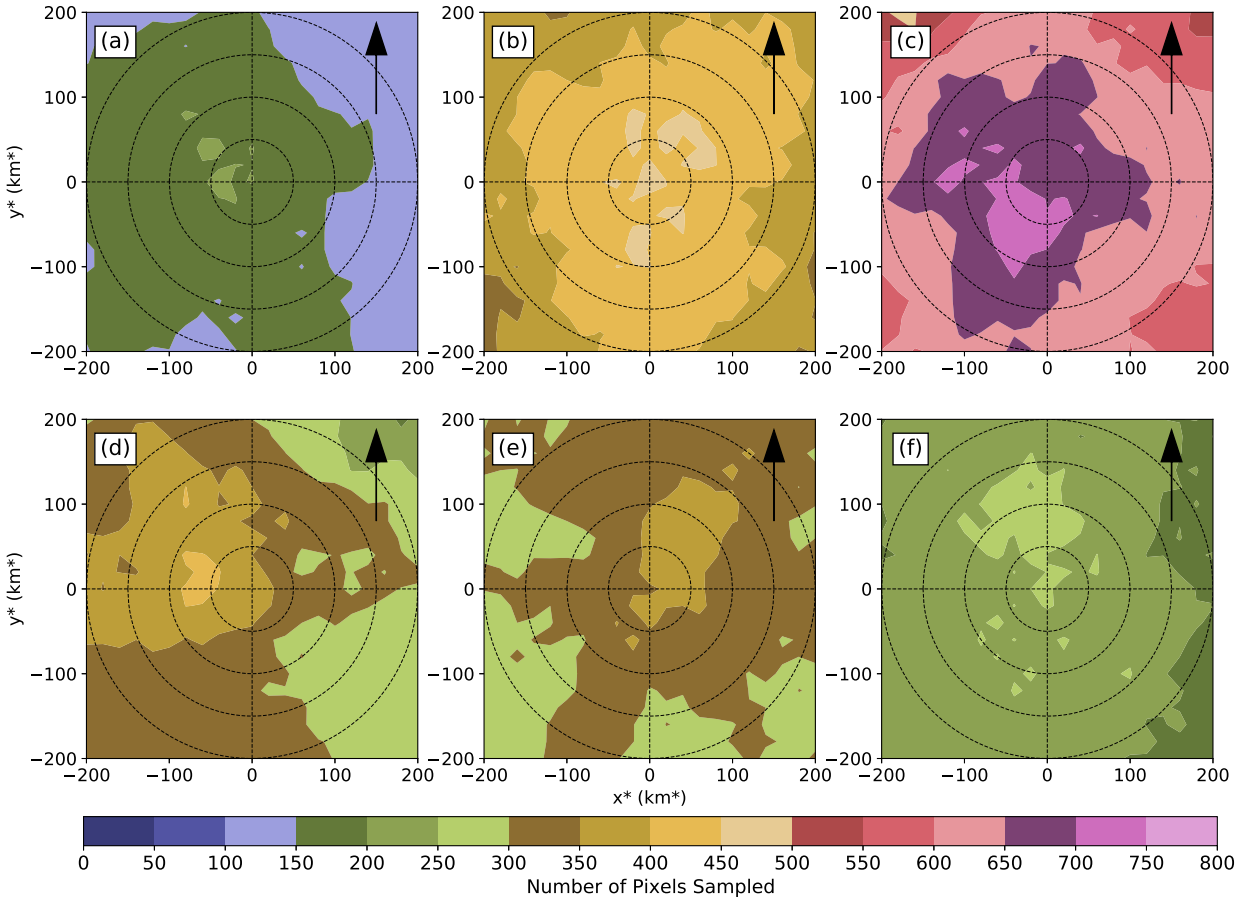

FIG. 4.2. Number of satellite pixels sampled in each scaled 20 km* × 20 km* bins for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical wind shear vector.

observation. However, the mean weakening rate steepened in the 12 hours following observation (from −3.89 knots to −14.44 knots; 1 kt = 0.514 m s^{-1}). Minor SS tropical cyclones were slowly intensifying in the 12 hours prior to the observation, but, by definition, plateau in intensity in the 12 hours following observation (5.19 kts to 0.88 kts). Minor IN tropical cyclones were already intensifying in the 12 hours prior to the observation and intensified more rapidly in the 12 hours following observation. In the minor IN group, 88 tropical cyclones had a 12-hour intensification rate of 15 knots or more. If we scaled the intensification rate to 24 hours, these 88 tropical cyclones would meet the category for rapid intensification (30 knots in 24 hours).

Like minor WK and SS tropical cyclones, major WK and SS tropical cyclones have similar mean intensities, while major IN tropical cyclones have a lower mean intensity at the time of the observation. On average, major WK tropical cyclones were slowly intensifying in the 12 hours prior to observation, and were weakening significantly in the 12 hours following observation (3.23 kts to −14.46 kts). Meanwhile, major SS tropical cyclones were slowly intensifying in the 12 hours prior to observation, but remained steady in intensity in the 12 hours following observation (8.93 kts to −0.16 kts) — again, by definition. Major IN tropical cyclones were intensifying in the 12 hours prior to observation, but intensified at a weaker rate in the 12 hours following observation. The decrease in the intensification rate for major IN tropical cyclones between the 12 hours prior and the 12 hours following observation indicates that they were reaching closer to their maximum potential intensity. Unsurprisingly, major tropical cyclones have a much smaller margin to further intensify.

The environment at the time of the observation is similar across the intensity change groups. While the mean sea surface temperature (SST) increases very slightly from WK to SS and IN for both minor or major tropical cyclones, the mean SST across all intensity change groups is over 28°C. The environmental vertical shear decreases from WK to SS and IN for both minor and major tropical cyclones. Major IN tropical cyclones have the most favorable vertical shear magnitude for tropical cyclone intensification or sustenance, with a mean of 4.18 m s^{-1} and a standard deviation of 2.0 m s^{-1}. Conversely, minor WK tropical cyclones have the least favorable shear magnitude for tropical cyclone intensification or sustenance, with a mean of 6.1 m s^{-1}. Similarly, the mean 850–500 mb environmental relative humidity trends more favorably for tropical cyclone intensification or sustenance from WK to SS and IN. But the mean 850–500 mb environmental relative humidity occurs over a relatively small range between 73% and 77%.

Overall, while the environment trends towards more favorable conditions from WK to SS and IN for both minor and major tropical cyclones, their mean magnitudes are similar. Therefore, one can infer that the difference in the distribution of convection and precipitation in our

composite analysis contribute towards the observed 12-hour future intensity change. However, in addition to being mindful about the smaller sample size in minor WK tropical cyclones, we must also be mindful of the intensity change in the 12 hours prior to the observation. As we have shown for our analysis, weakening tropical cyclones were already weakening and intensifying tropical cyclones were already intensifying in the 12 hours prior to the observation. Any signal in the distribution of convection and precipitation may be the cause or the symptom of the intensity change. Nonetheless, with this nuance in mind, we can still obtain useful information about the distribution of convection and precipitation in tropical cyclone intensity change from our composite analysis.

4.3.2 Distribution of Rainfall Occurrence

Figure 4.3 shows the fractional occurrence of rainfall (stratiform, convective, other, and shallow) in the different intensity change groups. For our discussions hereafter, the upshear right (UR), downshear right (DR), downshear left (DL), and upshear left (UL) quadrant refers to the bottom right, top right, top left, and bottom left sections of each composite. Minor WK tropical cyclones have a region of high rainfall occurrence in the right-of-shear quadrants, increasing from UR to a maximum in DR (Fig. 4.3a). The left-of-shear quadrants have a smaller rainfall occurrence, with an overall minimum in the upshear (or downwind) portions of UL.

For minor SS tropical cyclones, rainfall occurrence increases from UR to a maximum in DL (Fig. 4.3b). From DL to UL, rainfall occurrence decreases again, reaching an overall minimum in UR. Minor IN tropical cyclones have the largest rainfall occurrence out of all minor tropical cyclones (Fig. 4.3)c. Rainfall occurrence of over 95% wraps around the left-of-shear quadrants of minor IN tropical cyclones, with pockets of similarly high rainfall occurrence in the right-of-shear quadrants. Overall, minor IN tropical cyclones also have a more axisymmetric distribution of rainfall occurrence, with over 85% rainfall occurrence wrapping all the way around the tropical cyclone center.

Rainfall occurrence in major WK tropical cyclone increases from UR to DL, with a maximum of over 95% occurrence in DL that extends beyond 100 km* radius (Fig. 4.3d). From DL to

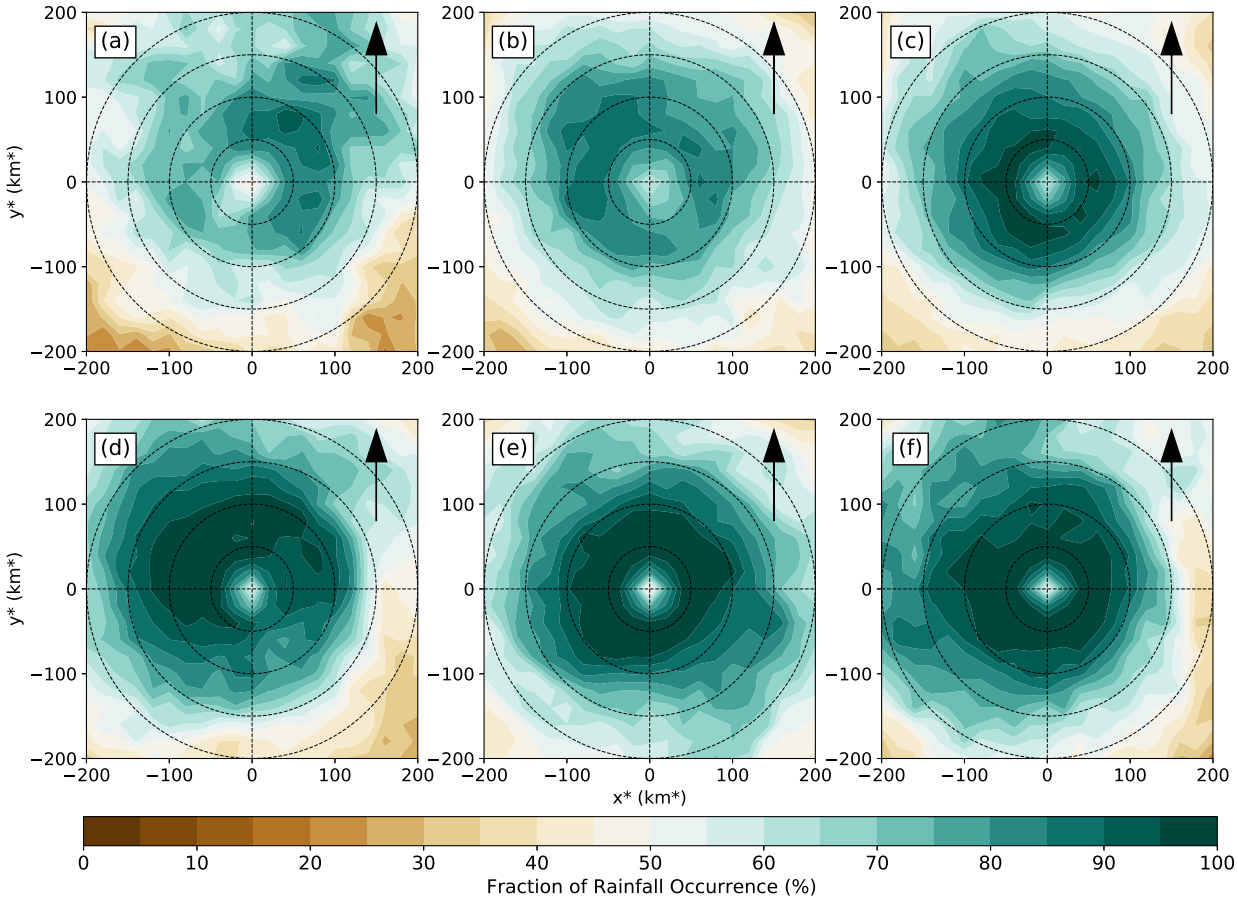

FIG. 4.3. Fractional occurrence of rainfall for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

UR, rainfall occurrence decreases to a minimum. For SS and IN major tropical cyclones, rainfall occurrence of over 95% wraps completely around the tropical cyclone center (Fig. 4.3e,f). Within 100 km* radius, the difference in rainfall occurrence between major SS and major IN tropical cyclones is not pronounced in this figure.

Looking at the differences in rainfall occurrence (Fig. 4.4), minor WK tropical cyclones have a smaller occurrence of rainfall closer to the center and in the left-of-shear quadrants, relative to the other minor tropical cyclones (Fig. 4.4a). But minor WK tropical cyclones also have a noticeably higher rainfall occurrence beyond 150 km* in the downshear quadrants, indicating the influence of stronger vertical shear. The difference in rainfall occurrence for

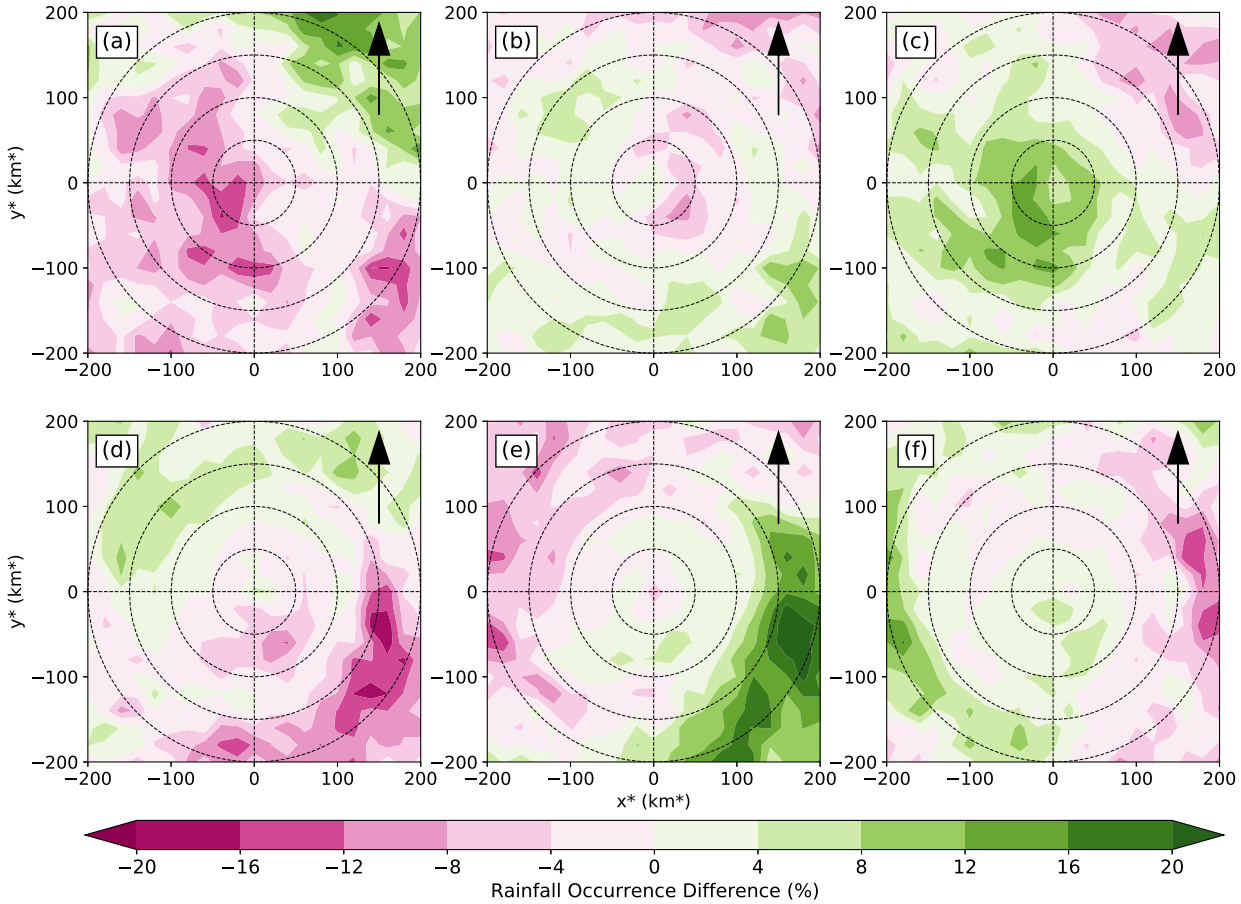

FIG. 4.4. Difference in fractional rainfall occurrence for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The difference is relative to the mean of their respective intensity category (i.e. minor and major). The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

minor SS tropical cyclones are generally smaller in magnitude (Fig. 4.4b). These smaller magnitudes indicate that in the context of rainfall occurrence, minor WK and minor IN tropical cyclones are on the opposite ends of the spectrum, while minor SS storms are in the middle. Nonetheless, there is a region of reduced rainfall occurrence in UR around 50 km * radius. Minor IN tropical cyclones contrast very differently to minor WK and SS tropical cyclones, possessing a large region of higher rainfall occurrence close to the center in all quadrants (Fig. 4.4c). This region of higher rainfall occurrence extends outwards up to 100 km* and 150 km* radius in DL and UL, respectively.

For major tropical cyclones, the differences in rainfall occurrence are small within ~150 km* radius. There is an area of higher rainfall occurrence in the downshear quadrants of major WK tropical cyclones, again hinting at the influence of stronger vertical shear (Fig. 4.4d). Close to the center, minor WK tropical cyclones have lower rainfall occurrence in UR. Major SS tropical cyclones have an area of significantly higher rainfall occurrence, mainly beyond 150 km in UR, which is also noticeable in the fraction of rainfall occurrence (Figs. 4.3e, Fig. 4.4e). The reason behind this higher rainfall occurrence is unclear. Recall that we have scaled our radius to account for different storm sizes. While the radius scaling is not perfect, this region contains observations from over 40 tropical cyclones (Fig. 4.1e) and includes over 250–300 pixels (Fig. 4.2e). This area of higher rainfall occurrence could indicate persistent rainband activity or precipitation symmetry that extends out to a larger radius. Since this is not the focus of our study, we will reserve further investigation for future studies.

Within 50 km* radius, major SS tropical cyclones have a slightly reduced rainfall occurrence. While between 50 km* and 100 km*, major SS have a slightly increased rainfall occurrence. In contrast, major IN tropical cyclones have a slight increase in rainfall occurrence from the center out to 100 km* radius in the upshear quadrants, and from the center out to 200 km* radius in UL (Fig. 4.4f).

Generally, the distribution of rainfall occurrence is consistent with previous studies. Rainfall occurrence maximizes in DL and becomes more axisymmetric as the intensification rate increases (from WK to SS and IN). Unlike previous studies, the rainfall occurrence maximizes in DR for minor WK tropical cyclones. The downshear distribution of the maximum rainfall occurrence indicates the influence of shear, while the lack of higher rainfall occurrence in the downwind quadrants (i.e., DL and UL) may be the result of the weakening tropical cyclone primary circulation. Since our sample size is the smallest in this intensity change category, we cannot assign too much significance on this pattern.

To summarize this section, minor IN tropical cyclones have increased rainfall occurrence and symmetry relative to other minor tropical cyclones. Recall that minor IN tropical cyclones have a weaker average intensity (Table 4.1). Since weaker tropical cyclones usually have a less

organized precipitation structure, the more symmetric rainfall occurrence in minor IN tropical cyclones is significant and is indicative of a necessary configuration of rainfall for intensification to occur. Conversely, rainfall symmetry alone is not enough for major tropical cyclones to intensify. While major WK tropical cyclones have a reduced rainfall occurrence in the upshear quadrants, the similarities in rainfall occurrence close to the center of major SS and IN tropical cyclones indicate that other factors must also play an influence on intensification.

4.3.3 Distribution of Convective and Stratiform Precipitation

We now investigate the distribution of convective and stratiform precipitation in intensity change. Figure 4.5 shows the convective fraction for all intensity change groups. Unlike the distribution of rainfall occurrence, the tropical cyclone eye is not apparent in Figure 4.5 because we calculate the convective fraction relative to raining pixels only, and that a majority of the raining pixels close to the center is dominantly convective. In addition, both convective and stratiform precipitation dominate the precipitation field in our composites. From the distribution of stratiform fraction (Fig. 4.6), we see that wherever the convective fraction is low (Fig. 4.5), the stratiform fraction is high.

Minor WK tropical cyclones have the highest convective fraction out of all minor tropical cyclones, with a maximum convective fraction in the downshear quadrants and close to the center (Fig. 4.5a). There are also local maxima in convective fraction beyond 150 km* radius. These local maxima may be deeper convection in the outer rainbands of a few tropical cyclones that dominate the small sample size in this intensity change group. Meanwhile, areas of larger stratiform fraction are dominant beyond 50 km* radius in the right-of-shear quadrants (Fig. 4.6a). In minor SS tropical cyclones, convective fraction exceeds 40% directly upshear of the center, with a large region of over 30% convective fraction extending from the center to about 50 km* radius in all quadrants (Fig. 4.5b). The stratiform fraction in minor SS is different than in minor WK tropical cyclones, with areas of larger stratiform fraction present beyond 100 km* and mainly in the downshear quadrants (Fig. 4.6b). Convective precipitation in minor IN tropical cyclones appears more organized, with a region of convective fraction over 30% rotating from DR to UL within 50 km* radius (Fig. 4.5c). However, the maximum

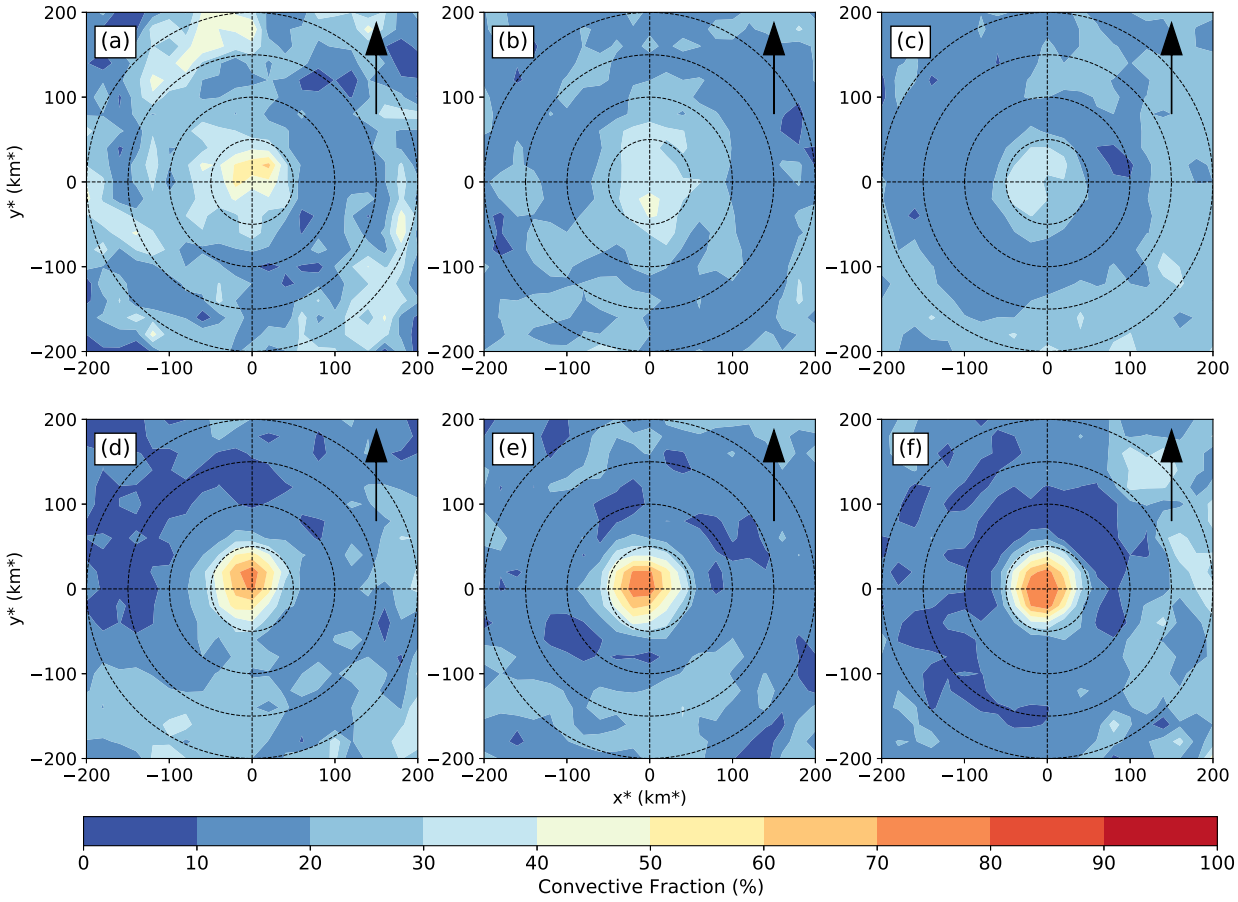

FIG. 4.5. Convective fraction for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

convective fraction in minor IN tropical cyclones is lower than in minor WK and SS tropical cyclones. In addition, stratiform precipitation also appears more organized, with an area of over 80% stratiform fraction occurring closer to the center and rotating from DR to UL (Fig. 4.6c).

The maximum convective fraction in all major tropical cyclones are between 70–80%. This higher convective fraction relative to minor tropical cyclones is simply a product of major tropical cyclones being more intense. But the size and location of the maximum convective fraction differs between major WK, SS, and IN. Major WK tropical cyclones have the smallest area of convective fraction over 70%, occurring in the downshear quadrants and within 50 km* radius (Fig. 4.5d). Simultaneously, maximum stratiform fraction occurs in DL beyond

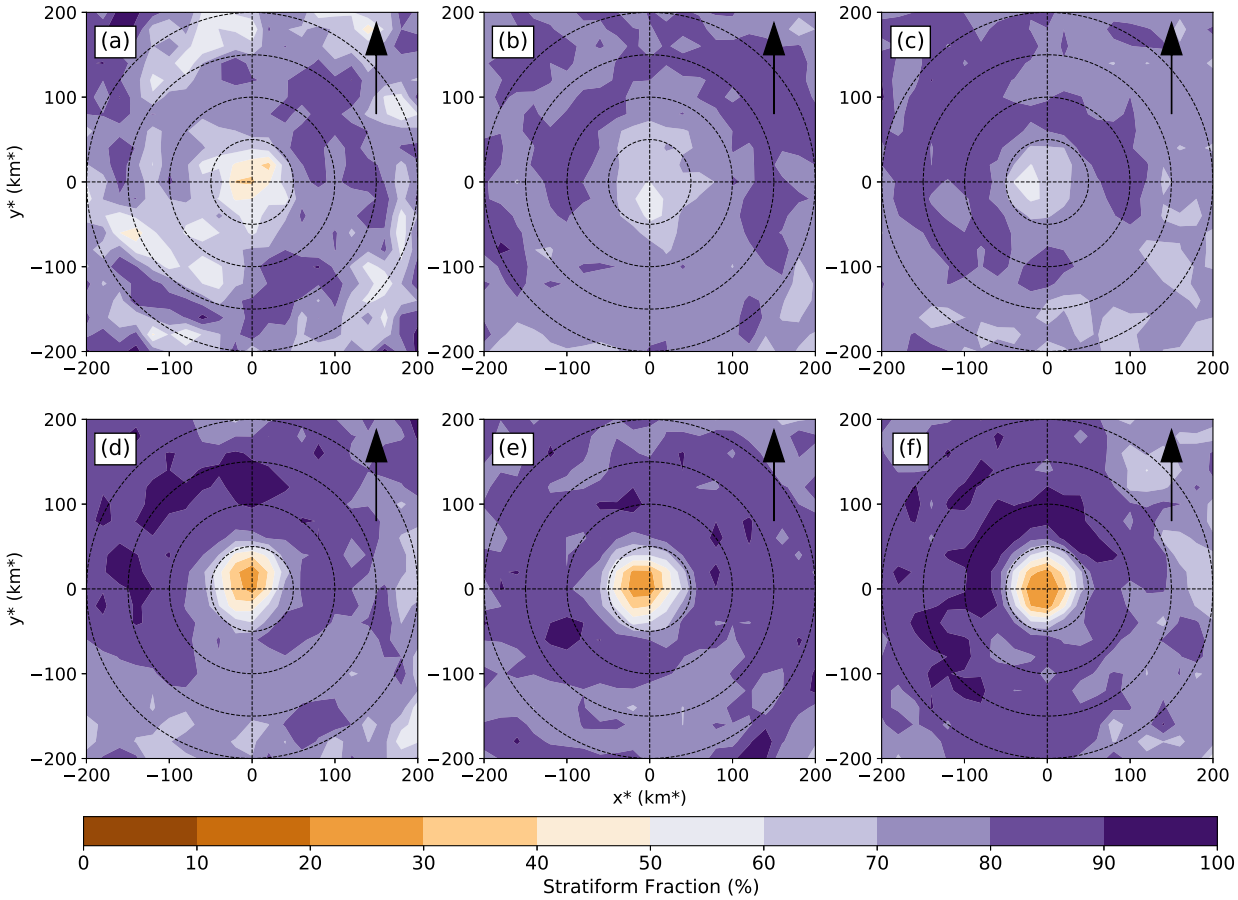

FIG. 4.6. Stratiform fraction for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

100 km* radius (Fig. 4.6d). The overall structure of stratiform precipitation is also more symmetric around the center of major WK storms, with a region of over 80% stratiform fraction wrapping nearly all the way around the center — breaking only in a small portion of UL. In major SS tropical cyclones, the area of convective fraction over 70% is larger relative to major WK storms, and extends azimuthally across DL within 50 km* radius (Fig. 4.5e). The stratiform precipitation is also more symmetric relative to major WK storms, with an area of over 80% stratiform fraction wrapping all the way around the tropical cyclone center (Fig. 4.6e). Finally, major IN tropical cyclones have the largest area of convective fraction over 70% that is most prevalent in the left-of-shear quadrants within 50 km* radius and extends into the upshear quadrants (Fig. 4.5f). Stratiform fraction in major IN tropical cyclones are also better

organized, with stratiform fraction over 90% rotating from DR to UL and stratiform fraction over 80% wrapping completely around the center (Fig. 4.6f).

To highlight the difference in convective fraction between the different intensity groups, we plot the convective fraction difference in Fig. 4.7. As we have shown, the pattern of stratiform fraction is essentially opposite to the pattern of convective fraction. Plotting the stratiform fraction difference produce a pattern that is opposite to the patterns in Fig. 4.7 (not shown). Therefore, while we discuss the convective fraction difference in terms of bins that have more or less convective fraction relative to the mean (positive or negative convective fraction difference, respectively), we can also interpret the locations as having more or less stratiform fraction (negative or positive convective fraction difference, respectively).

Relative to other minor tropical cyclones, minor WK tropical cyclones have more convective precipitation (positive convective fraction difference) in the downshear quadrants within 50 km* radius (Fig. 4.7a). Beyond 50 km* radius, minor WK tropical cyclones have more convective precipitation in DL and in the upwind portions of UL, which generally coincide with lower rainfall occurrence (Fig. 4.3a). In minor SS tropical cyclones, the convective fraction between 50 and 100 km* radius in the right-of-shear quadrants is higher than other minor tropical cyclones (Fig. 4.7b). Minor IN tropical cyclones display a significant difference in precipitation structure compared to minor WK and SS, with less convective precipitation within 100 km* radius in all quadrants and slightly more convective precipitation in the left-of-shear quadrants around 50 km radius (Fig. 4.7c).

For major tropical cyclones, the relationship between the distribution of convective precipitation with intensity change is relatively more straightforward. Major WK tropical cyclones have generally more convective precipitation between 50 and 100 km* radius in all quadrants, with less convective precipitation within 50 km* radius in the upshear quadrants (Fig. 4.7d). Convective fraction is also higher beyond 50 km* radius in UL. For major SS tropical cyclones, convective fraction is generally higher in DL beyond 50 km* radius (Fig. 4.7e). Finally, major IN tropical cyclones have a higher convective fraction in the upshear quadrants and within 50 km* radius (Fig. 4.7f). Between 50 and 100 km* radius, there is generally less convective

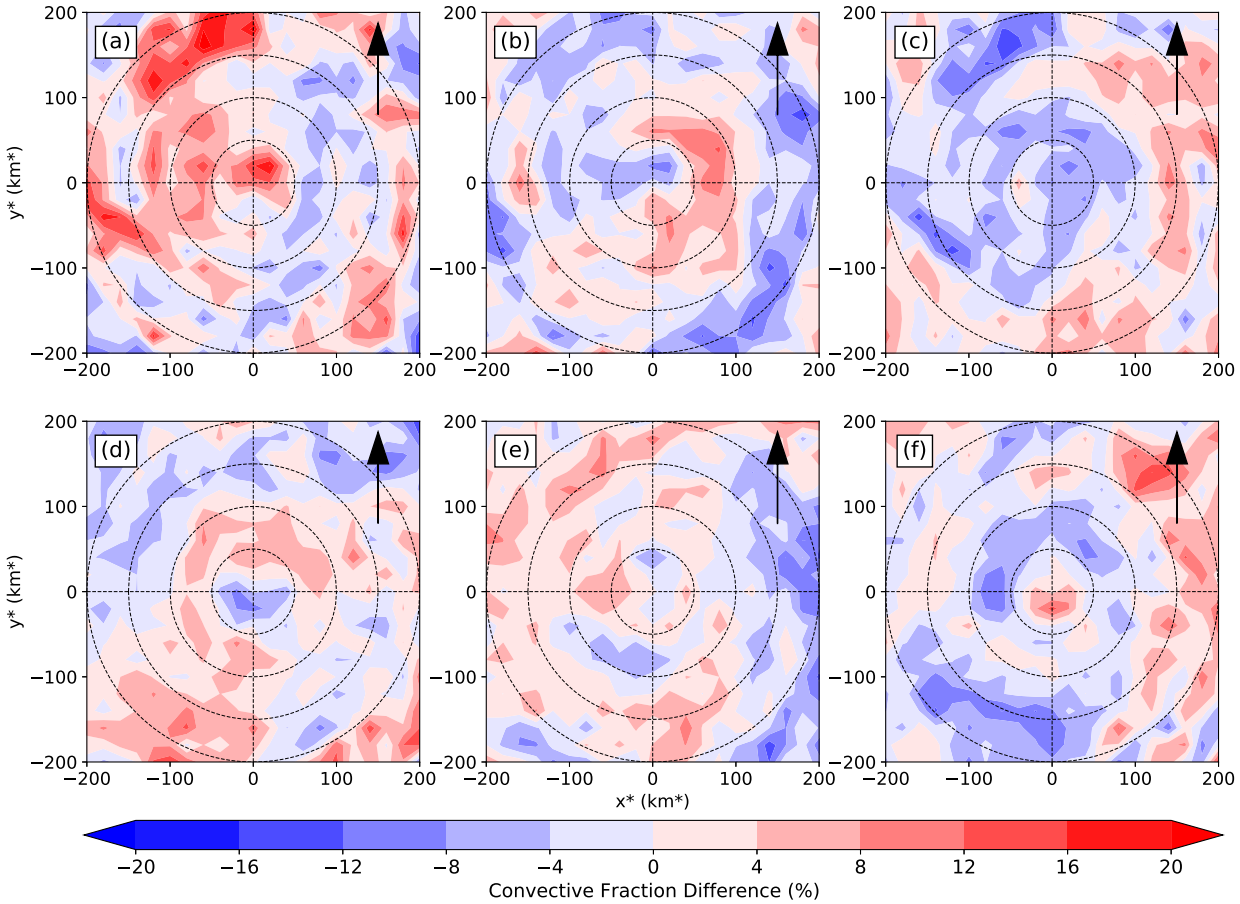

FIG. 4.7. Difference in convective fraction for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The difference is relative to the mean of their respective intensity category (i.e. minor and major). The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

precipitation, with pockets of marginally more convective precipitation in the upshear quadrants.

Relating the findings of this section to the findings of the previous section, minor IN tropical cyclones have a more symmetric precipitation structure that consists of overall less convective precipitation relative to other minor tropical cyclones. But in the left-of-shear quadrants closer to the center, there is marginally more convective precipitation relative to other minor tropical cyclones. Conversely, while major IN tropical cyclones have only marginally more precipitation occurrence than other major tropical cyclones, they have a region of higher convective fraction within the upshear quadrants and close to the center.

4.3.4 Distribution of Mean Precipitation Rate

Now that we have a better understanding of the distribution of convective precipitation in tropical cyclone intensity change, we investigate their contribution towards precipitation. Figure 4.8 shows the mean surface precipitation rate from GPROF. We averaged the GPROF surface precipitation rate only for the regions that the precipitation type classifier identified as raining. Averaging the raw GPROF surface precipitation rate without first separating raining and non-raining regions produce a similar distribution but with smaller magnitudes (not shown). Figure 4.9 shows the difference in surface precipitation rate relative to the mean of the each intensity group, while Figs. 4.10 and 4.11 show the mean precipitation rates for convective and stratiform precipitation, respectively. Note that while the mean convective precipitation is higher than the mean stratiform precipitation, we focus on their magnitudes and locations of local maxima relative to other intensity change groups.

Minor WK tropical cyclones have a disorganized distribution of mean precipitation rate, with local maxima generally in the downshear quadrants within 100 km* radius (Fig. 4.8a). Minor SS tropical cyclones have a band of higher mean precipitation rate that appears to spiral inward from 100 km* radius in DR into DL, before reaching the upshear quadrants at around 50 km* radius (Fig. 4.8b). In minor IN tropical cyclones, the mean precipitation rate distribution appears the most organized out of all minor tropical cyclones, with a maximum in the left-of-shear quadrants (Fig. 4.8c). The distribution of mean precipitation rate in major tropical cyclones shows a more systematic trend. In major WK tropical cyclones, high mean precipitation rate stretch from the downwind region of DR into DL (Fig. 4.8d). Relative to major WK tropical cyclones, high mean precipitation rates occur farther downwind in major SS while simultaneously expanding farther into UL (Fig. 4.8e). In major IN tropical cyclones, high mean precipitation rates engulf the entire left-of-shear quadrants close to the tropical cyclone center, including the regions immediately upshear and downshear of the center.

Looking at the precipitation rate difference, minor WK tropical cyclones have scattered regions of higher precipitation rate that roughly coincide with the local maxima in the mean surface precipitation rate (Fig. 4.8a), found at: i) along the y = 0 km* line within 25 km*, ii)

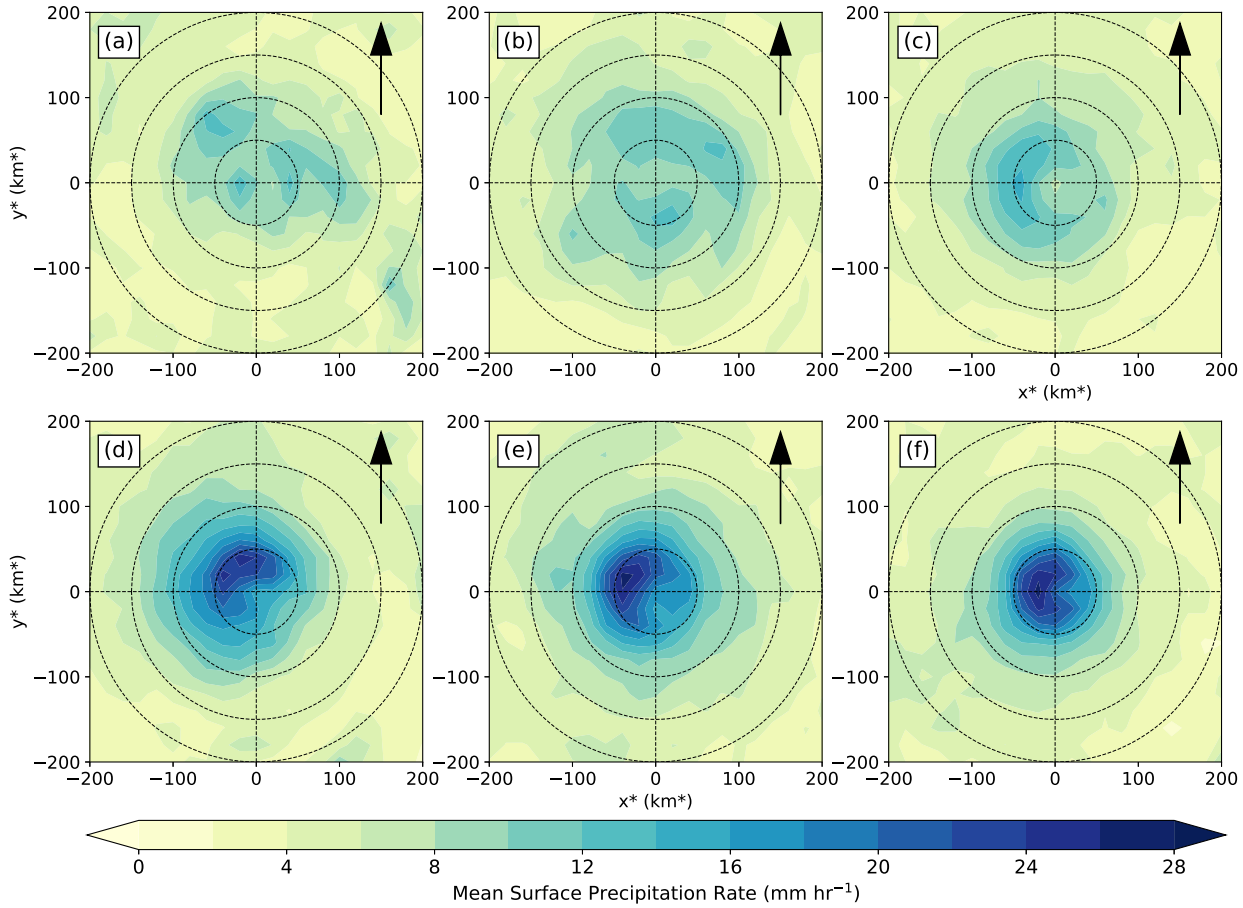

FIG. 4.8. Mean surface precipitation rates for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

along the y = 0 km* line between 50 and 150 km* radius extending larger radii upwind, and iii) around 100 km* radius in DL (Fig. 4.9a). The first region of higher mean precipitation rate coincides with higher convective precipitation rate (Fig. 4.10a). The second region of higher mean precipitation rate coincides with both higher convective and stratiform precipitation rates (Figs. 4.10a, 4.11a). The third region of higher mean precipitation rate coincides with higher stratiform precipitation rate (4.11a). Recall that beyond 150 km* radius in DL, high convective fraction dominates in an area with low rainfall occurrence. However, the total and convective precipitation rates there are generally low. More importantly, minor WK tropical cyclones have a much lower mean precipitation rate relative to other minor tropical cyclones in UL and immediately upshear of the center between 50 and 150 km* radius.

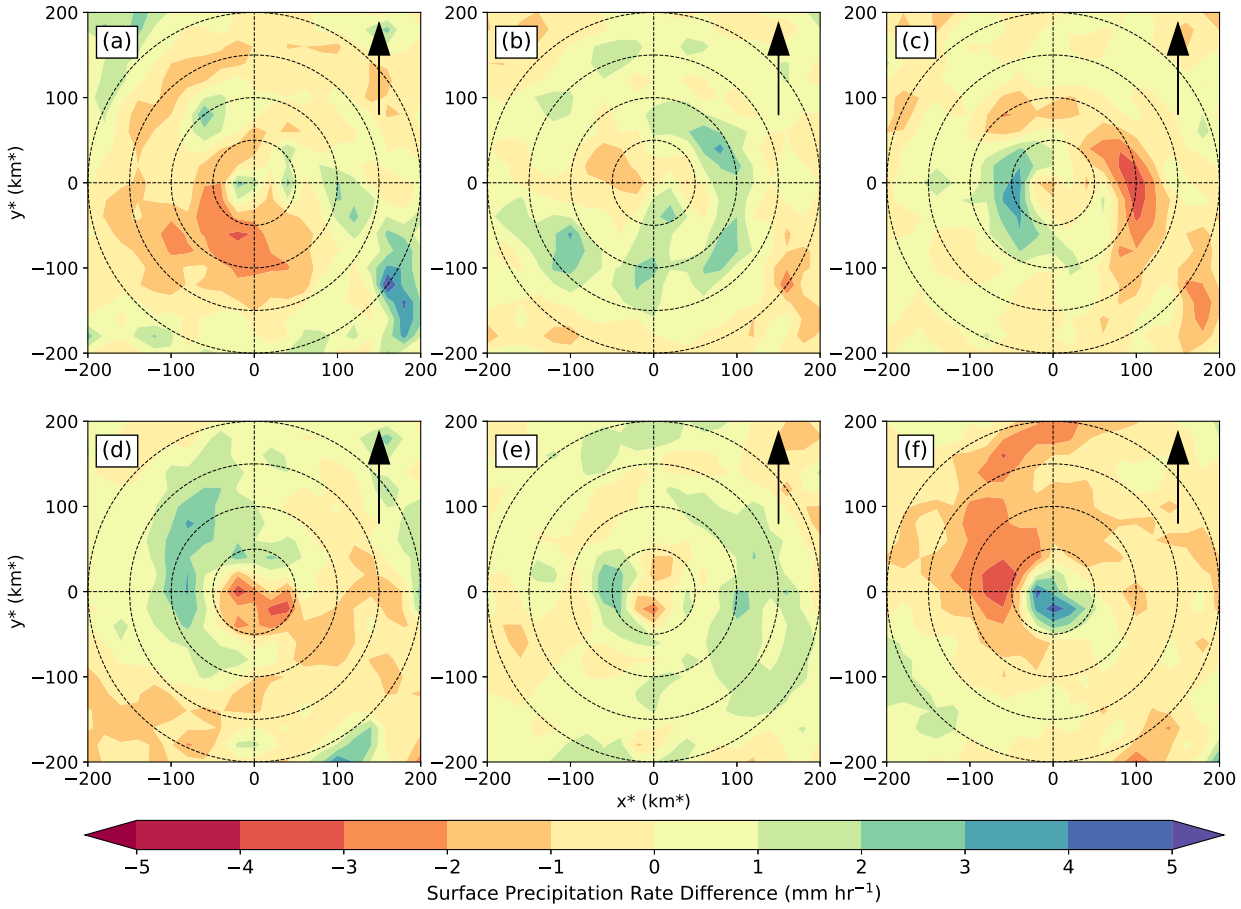

FIG. 4.9. Difference in mean surface precipitation rates for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The difference is relative to the mean of their respective intensity category (i.e. minor and major). The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

Minor SS tropical cyclones contain an arc of higher mean precipitation rate that extends from around 125 km* radius in UR to 50 km* radius in DR, with additional regions of localized maxima immediately upshear of the center and in UL (Fig. 4.9b). Higher convective and stratiform precipitation rates both contribute to these local maxima in mean precipitation rate (Figs. 4.10b and 4.11b). Conversely, minor IN tropical cyclones have a distinct arc of higher mean precipitation rate in the left-of-shear quadrants (Fig. 4.9c). This region coincides with both higher convective and stratiform precipitation rates (Figs. 4.10c and 4.11c). In addition, a coherent arc of higher stratiform precipitation also extends downwind of this precipitation maxima into the upshear quadrants.

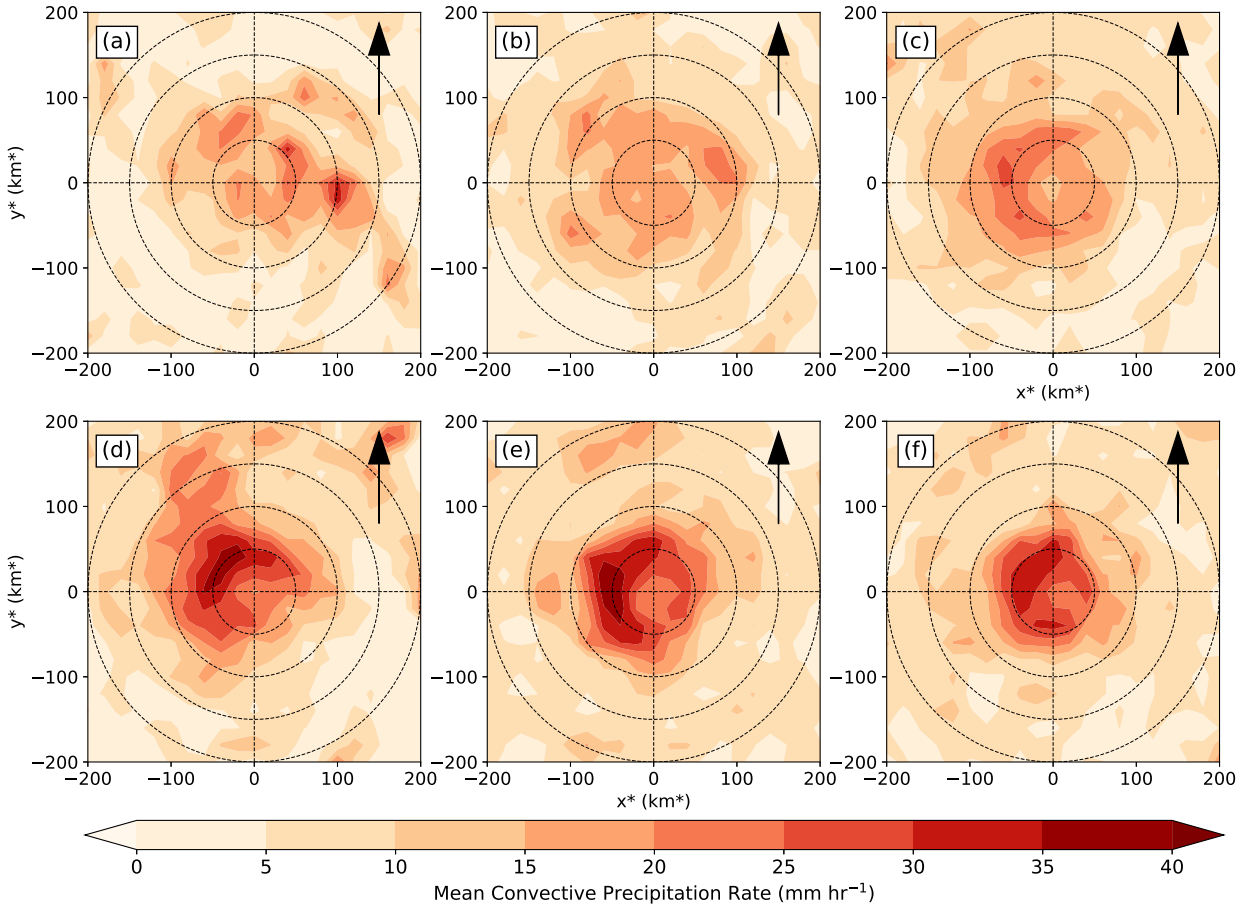

FIG. 4.10. Mean convective precipitation rates for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

The precipitation rate difference in major tropical cyclones (Fig. 4.9d,e,f) reflect the systematic trend present in the mean surface precipitation rate (Fig. 4.8d,e,f). That is, going from major WK to SS and IN, the precipitation rate difference becomes increasingly more negative in DL between 50 and 150 km* radius (Fig. 4.9d,e,f). In addition, major SS tropical cyclones have higher mean precipitation rate in the right-of-shear quadrants beyond 50 km* radius (Fig. 4.9d). More importantly, major IN tropical cyclones have a region of stronger mean precipitation rate in the upshear quadrants within 50 km* radius. Generally, both convective and stratiform precipitation contribute to this distribution of mean precipitation rate in major tropical cyclones (Fig. 4.10d,e,f and 4.11d,e,f). In summary, major WK tropical cyclones have higher mean precipitation rate beyond 50 km* in DL, major SS tropical cyclones have a

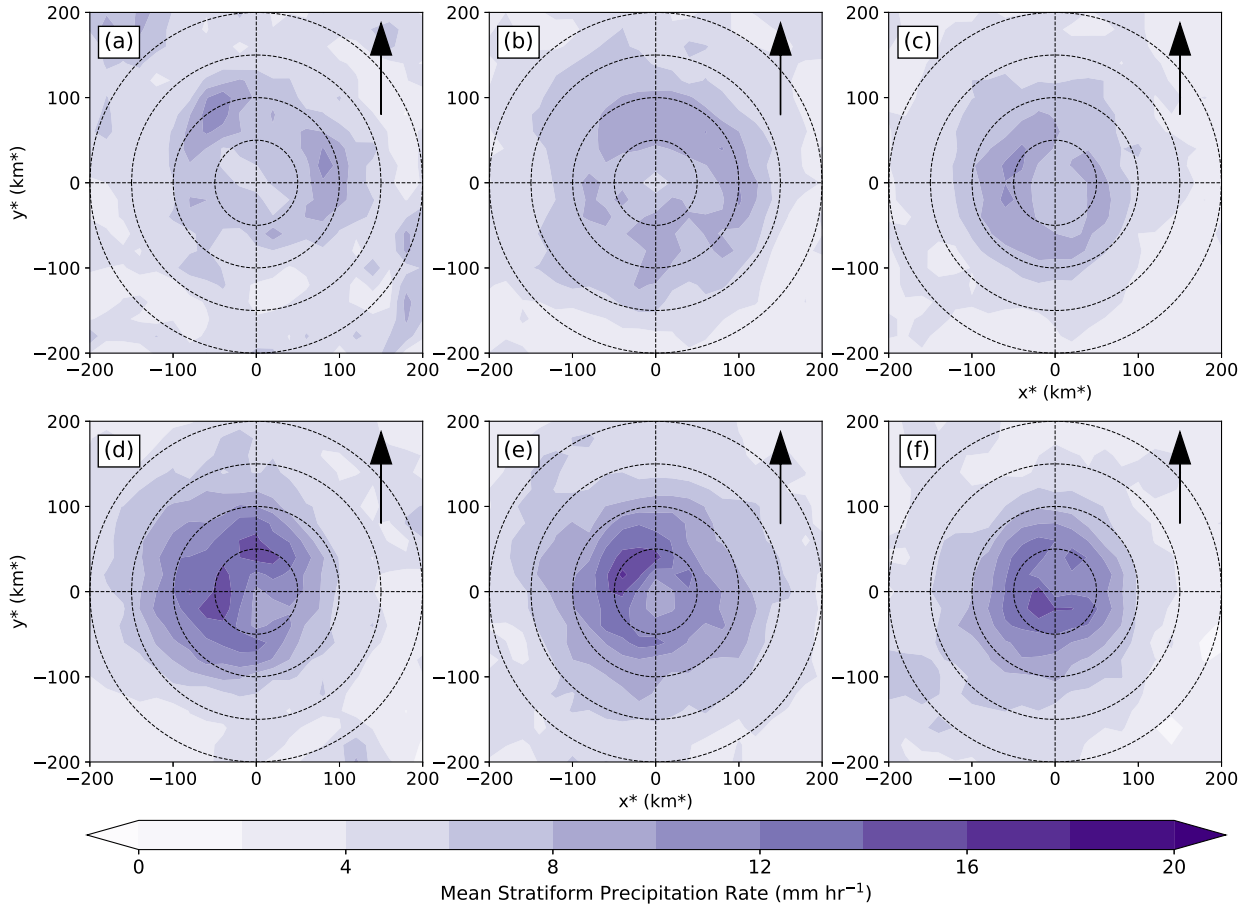

FIG. 4.11. Mean stratiform precipitation rates for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

more symmetric region of higher mean precipitation rate beyond 50 km* radius, and minor IN tropical cyclones have a higher mean surface precipitation rate closer to the tropical cyclone center and in the upshear quadrants.

Linking the findings of previous sections to this section, we find that minor WK and SS tropical cyclones tend to be more asymmetric, with a more disorganized distribution of convection and precipitation. In contrast, minor IN tropical cyclones tend to be more symmetric, contain less convective precipitation, and have a region of high precipitation rate in the left-of-shear quadrants. Precipitation occurs in a more symmetric fashion in all three major tropical cyclone intensity categories. Going from major WK to SS and IN tropical cyclones, the increasingly negative precipitation rate difference beyond 50 km* radius in DL indicate the reduced

influence of vertical shear as intensification rate increases, resulting in higher mean precipitation rate closer to the center. More importantly, major IN tropical cyclones have a convective fraction maximum in the upshear quadrants and close to the tropical cyclones center.

4.3.5 Intensity of Convection

Finally, we compare our convection and precipitation analyses with a different observation in the form of the 89.0-GHz PCT. Recall that at the 89.0-GHz frequency, precipitation-sized ice particles in deep convection scatters terrestrial radiation. Intense updrafts in really strong convection can keep these ice particles aloft for a longer period, enabling them to grow and scatter more terrestrial radiation, thus producing a much lower brightness temperature at this frequency. In this section, we use the coldest 5 percent (5th-percentile) of the 89.0-GHz PCT in each 20×20 km* bin as proxy for the intensity of convection (Fig. 4.12).

Generally, the 5th-percentile 89.0-GHz PCT field agrees with our previous analysis. Minor WK and SS tropical cyclones have scattered and isolated regions of deep convection (Fig. 4.12a,b). The coldest 5th-percentile 89.0-GHz PCT in minor SS tropical cyclones are comparable to the coldest 5th-percentile 89.0-GHz PCT in minor IN tropical cyclones, indicating that the intensity of convection between the two groups are similar (Fig. 4.12b,c). More importantly, the biggest distinction among the minor tropical cyclones is that minor IN tropical cyclones have a more symmetric distribution of moderate 5th-percentile brightness temperature (~200 K). When combined with our convective and precipitation analyses, the 5th-percentile 89.0-GHz PCT field indicates that the more symmetric precipitation in minor IN tropical cyclones consist of mostly moderate convection.

Among major tropical cyclones, major IN tropical cyclones have distinctly more intense convection occurring over a larger sector of the tropical cyclone inner core (within 50 km*), particularly in the upshear quadrants (Fig. 4.12f). In contrast, the most intense convection in major WK and SS tropical cyclones are not as intense and remain in the left-of-shear and DL, respectively (Fig. 4.12d,e). The 5th-percentile 89.0-GHz PCT field indicates that the increased convective fraction in the upshear quadrants of major IN tropical cyclones coincide with increased convective intensity there.

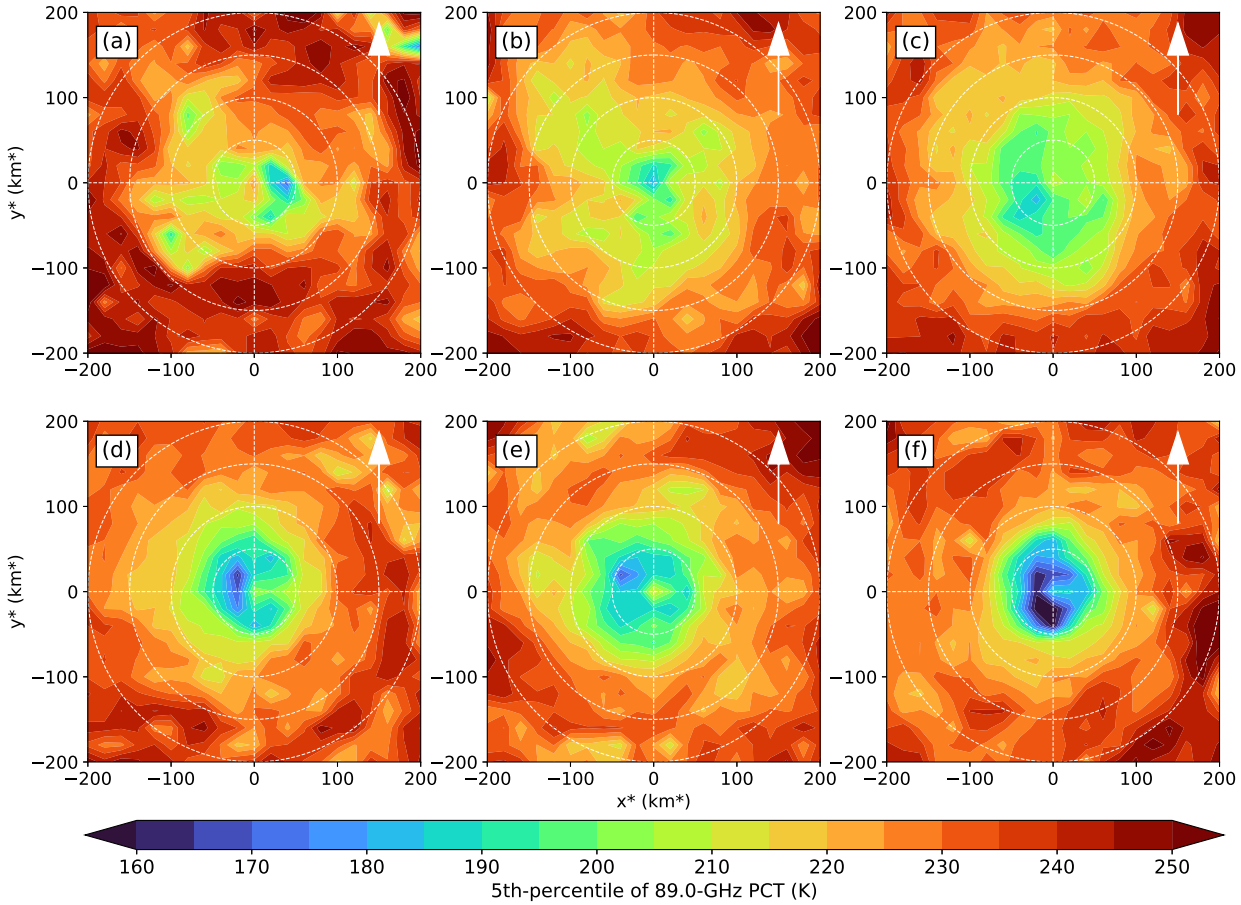

FIG. 4.12. Fifth-percentile 89.0-GHz PCT for (a) minor WK, (b) minor SS, (c) minor IN, (d) major WK, (e) major SS, and (f) major IN intensity change group. The range rings are every 50 km* and the arrow in the top right of each panel shows the environmental vertical shear vector.

4.4 DISCUSSION

In our composite analysis, the intensifying tropical cyclones were already intensifying and the weakening tropical cyclones were already weakening at the time of the GPM observations. While the magnitude of the mean intensity change (weakening or intensifying) generally increased in the 12 hours following the GPM observation, we cannot explicitly attribute our findings here as being the cause of the subsequent intensity change. Nonetheless, we can discuss our findings in the context of the configurations of convection and precipitation that are important or necessary for intensification.

From our analysis, a more symmetric rainfall coverage is more important for the intensification of minor tropical cyclones. The more symmetric rainfall coverage consists of less convective precipitation relative to non-intensifying minor tropical cyclones, with moderate convective intensity. Convection of similar intensity or stronger are present in non-intensifying minor tropical cyclones, but are more randomly distributed in pockets of high convective fraction within regions of lower rainfall occurrence. The resulting mean precipitation rate is more scattered for non-intensifying tropical cyclones, but have a left-of-shear maximum in intensifying minor tropical cyclones. In addition, increased convective and stratiform precipitation rates wrap around the upshear quadrants more coherently. We hypothesize that for minor tropical cyclones to intensify, latent heat release from the more symmetric distribution of precipitation is more efficient at spinning up the tropical cyclone compared to asymmetric and randomly-distributed deep convection.

For major tropical cyclones, rainfall occurrence within ∼100 km* radius is already more symmetric across all intensity change groups. However, non-intensifying major tropical cyclones have i) higher convective fraction and mean precipitation rate occurring at a larger radii; and ii) weaker convective intensities confined to the left-of-shear quadrants. Conversely, intensifying major tropical cyclones have i) higher convective fraction and mean precipitation rate occurring closer to the tropical cyclone center, particularly in the upshear quadrants; and ii) more intense convection that wraps around the left-of-shear quadrants into the upshear quadrants. Therefore, we hypothesize that for major tropical cyclones to intensify, this configuration of convection and precipitation is optimal for the conversion of latent heating to kinetic energy.

4.5 Conclusion

Using observations from the GPM satellite, we have analyzed the distribution of convection and precipitation in tropical cyclone intensity change. The GPM data comes from the Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED; Razin et al. 2021). We obtained the distribution of convection and precipitation using a precipitation type classifier for the GPM satellite that we developed in Chapter 3. Using this

precipitation type classifier, we were able to leverage the wider swath of the GMI to simultaneously observe a large portion of the tropical cyclone and increase our composite sample size. We grouped our tropical cyclone observations based on their future 12-hour intensity change and their intensity at the time of the GPM observation. In addition, we minimized the influence of the environmental variability by selecting overpasses of tropical cyclones located in favorable environments that include sea-surface temperatures $\geq 27°C$ and vertical shear magnitude ≤ 10 m s^{-1}. The tropical cyclones must also remain over 150 km from land in the following 12 hours.

In our composite analysis, the intensifying tropical cyclones were already intensifying and the weakening tropical cyclones were already weakening prior to the GPM observation. Therefore, we cannot explicitly infer that the distribution of convection and precipitation in our composites causes the intensity change in the period following the GPM observation. Rather, we can interpret the distribution of precipitation as being important for the observed intensity change.

We found that intensifying minor tropical cyclones have a more axisymmetric distribution of rainfall that have less convective precipitation relative to other minor tropical cyclones and generally moderate convective intensity. The precipitation rate distribution is also more coherently encircling the tropical cyclone center. We hypothesize that a more symmetric distribution of precipitation and convection of moderate intensity is more important for minor tropical cyclones to intensify, compared to more scattered and randomly organized convection. Conversely, rainfall occurrence in major tropical cyclones are already more symmetric. What sets intensifying major tropical cyclones apart from non-intensifying major tropical cyclones is the occurrence of high convective fraction and mean precipitation in the upshear quadrants and closer to the center, coinciding with a more intense convection. We hypothesize that the optimal location of deep convection and precipitation for the intensification of major tropical cyclones is in the upshear quadrants close to the tropical cyclone center.

Using the GMI-based precipitation type classifier that we developed in Chapter 3 enabled us to obtain a larger sample size of precipitation type observation compared to if we were to

use the precipitation type observation from the DPR. But we still faced issues with unbalanced sample sizes, with minor weakening tropical cyclones having the smallest sample size. We would benefit from a similar precipitation type classifier for other passive microwave sensors to increase our sample size, which would enable us to adopt the following improvements to our analysis:

- further separate the observations to account for the intensity change prior to the GPM observation
- further separate intensifying tropical cyclones into rapidly intensifying (RI) and non-RI tropical cyclones to see if our current hypothesis still holds, especially since RI tropical cyclones (when scaled to 24-hour intensity change) dominate our sample for the minor IN group
- average the precipitation fraction in batches of observations within each intensity change group to reduce the noise in the current approach

Since our precipitation type classifier uses the passive microwave observations that have relatively big instantaneous field-of-view, we cannot directly observe or infer the fraction of shallow convection and their relationship to tropical cyclone intensity change. Nonetheless, we observed distinct signals in both intensifying minor and major tropical cyclones. Therefore, we must attempt to gain a better understanding of the factors that contribute to the orientation of convection and precipitation observed in this analysis in order to develop forecast applications.

CHAPTER 5

CONCLUSIONS

In an effort to analyze the distribution of precipitation in tropical cyclone intensity change, this dissertation has accomplished three primary objectives. The first accomplished objective is the compilation of an extensive research-quality and open-access dataset of low-Earth orbit (LEO) satellite passive microwave observations centered on tropical cyclones, called the Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED). TC PRIMED consists of tropical cyclone-centric 1) inter-calibrated, multi-channel, multi-imager microwave brightness temperatures, 2) retrieved rainfall from NASA's Goddard Profiling algorithm (GPROF), 3) nearly coincident geostationary satellite infrared imagery, and 4) auxiliary data such as tropical cyclone position and intensity, ERA5 fields and derived environmental diagnostics, and satellite precipitation radar variables. TC PRIMED includes observations from over 168,000 LEO satellite overpasses of 2,101 tropical cyclones from 1998 to 2019. A simple composite analysis of retrieved rainfall, passive microwave brightness temperatures, and infrared brightness temperatures relative to the tropical cyclone lifetime maximum intensity demonstrated the huge potential for TC PRIMED and how the dataset provides the scientific community access to a large, long-term, global, research-quality, and open-access dataset of LEO satellite observations of tropical cyclones with associated infrared and environmental data.

Work is ongoing to publish TC PRIMED on a permanent and public online repository. Future work on TC PRIMED involves:

- updating TC PRIMED to include observations from the 2020 tropical cyclone season to the most recent tropical cyclone season
- adding other variables to the dataset such as the vertical profile of derived latent heating from GPROF and a land/ocean flag for the passive microwave observations (which are already available for the GPROF and precipitation radar variables from their respective original files)

- adding satellite scan and orientation information to allow users to correct for any parallax errors
- adding observations from new satellites or constellations such as the Time-Resolved Observations of Precipitation structure and storm Intensity with a Constellation of Smallsats (TROPICS)

The second accomplished objective is the successful application of the random forest model to classify precipitation types — such as convective and stratiform precipitation — from the GMI. The steps taken to accomplish this objective involve i) obtaining GPM satellite observations of tropical cyclones from TC PRIMED, ii) generating input predictors from the GMI observations, which include raw brightness temperatures, derived brightness temperatures and brightness temperature texture information, and iii) training the random forest model to classify the precipitation type using the GPM DPR observation as reference. The model performs very well at delineating heavily precipitating regions from non-raining regions and captures key features like the tropical cyclone eye, eyewall, and primary rainbands. But the model struggles to classify randomly distributed and scattered convective precipitation, resulting in the detection rate for no rain, stratiform, and convective precipitation type of over 96%, 80% and 50%, respectively.

Analysis of the model's prediction probabilities for correct and incorrect classifications reveals that the model is overconfident when it incorrectly classified convective and stratiform precipitation (e.g., convective precipitation that the model classified as no rain or stratiform). More importantly, a drop-column analysis reveals that the most important predictor is texture information. Certain drawbacks in this precipitation classification approach warrants further improvement, such as:

- employing a different method to better represent the various precipitation types within the GMI IFOV
- organizing the training dataset such that each training, calibrating, and testing sample are completely independent of each other to reduce model over-fitting

- experimenting with other random forest hyperparameters to improve model performance and reduce over-fitting
- employing a convolutional neural network model instead of the random forest to better represent the brightness temperature texture information
- extending this approach to include other passive microwave sensors such as the Tropical Rainfall Measurement Mission, which has its own precipitation radar observation. This approach is also possible for other passive microwave sensors without their own precipitation radar, like the Advanced Microwave Scanning Radiometer, which has similar passive microwave frequencies and corresponding IFOV as the GMI

The ultimate objective of this dissertation is to analyze the distribution of convection and precipitation in tropical cyclone intensity change using both TC PRIMED and the precipitation type classification model. The analysis includes GPM observations of tropical cyclones in favorable environments such as sea-surface temperatures of over 27°C and vertical shear magnitude of under 10 m s^{-1}. The observed tropical cyclone must also remain at least 150 km away from land throughout the 12-hour period following observation. The observations were separated into six intensity change groups based on their future 12-hour intensity change and intensity at the time of the observation. The groups are weakening, steady-state, and intensifying minor and major tropical cyclones. Subsequently, shear-relative composites of rainfall occurrence, convective/stratiform fraction, precipitation rate, and 5th-percentile of the 89.0-GHz PCT were generated. Improvements to the current approach requires the development of the precipitation type classification algorithm for other passive microwave sensors in order to increase the sample size and i) account for a more diverse range of intensity change rate, ii) account for intensity changes prior to observation, and iii) obtain a smoother analysis of precipitation fraction.

Results reveal that intensifying minor tropical cyclones have: i) a more symmetric distribution of rainfall, ii) lower convective fraction overall compared to non-intensifying minor tropical cyclones, and iii) convection that are of moderate intensity. Rainfall occurrence in major tropical cyclones are generally more axisymmetric. However, relative to non-intensifying

major tropical cyclones, intensifying major tropical cyclones have more numerous and more intense deep convection with increased precipitation rate in the upshear quadrants, closer to the tropical cyclone center. The results lead to the following hypotheses: i) for minor tropical cyclones to intensify, an axisymmetric distribution of precipitation that are not necessarily convective is more important compared to more randomly distributed convection, and ii) for major tropical cyclones to intensify, the occurrence of intense deep convection and increased precipitation in the upshear quadrants are optimal for the efficient conversion of latent heating into kinetic energy.

While the convection and precipitation signals in tropical cyclone intensity change are robust, the analysis would benefit from a larger sample size of passive microwave observations that will allow for:

- the consideration of intensity changes prior to the observation and/or onset of intensification, thus providing a better understanding of the precursors to intensification
- further separation of intensity change categories to account for rapidly intensifying tropical cyclones
- the reduction of noise compared to the current approach

This dissertation advances the science by i) providing the scientific community with an extensive dataset of satellite passive microwave observations of tropical cyclones with ancillary data, ii) demonstrating the utility of machine learning in distinguishing precipitation types from passive microwave observations while providing guidelines for future improvements to such approach, and iii) clarifying the role of convection and precipitation in tropical cyclone intensity change. The findings herein can guide the development of or improvements to tropical cyclone intensity forecast products.

REFERENCES

Alvey, G. R., III, J. Zawislak, and E. Zipser, 2015: Precipitation properties observed during tropical cyclone intensity change. *Mon. Wea. Rev.*, **143**, 4476–4492, doi: 10.1175/MWR-D-15-0065.1.

Anagnostou, E. N. and C. Kummerow, 1997: Stratiform and convective classification of rainfall using SSM/I 86-GHz brightness temperature observations. *J. Atmos. Oceanic Technol.*, **14**, 570–575.

Awaka, J., M. Le, V. Chandrasekar, N. Yoshida, T. Higashiuwatoko, T. Kubota, and T. Iguchi, 2016: Rain type classification algorithm module for GPM dual-frequency precipitation radar. *J. Atmos. Oceanic Technol.*, **33**, 1887–1898, doi: 10.1175/JTECH-D-16-0016.1.

Backus, G. and F. Gilbert, 1970: Uniqueness in the inversion of inaccurate gross earth data. *Philos. Trans. Roy. Soc. London*, **266**, 123–192.

Berg, W., et al., 2016: Intercalibration of the GPM microwave radiometer constellation. *J. Atmos. Oceanic Technol.*, **33**, 2639–2654, doi: 10.1175/JTECH-D-16-0100.1.

Brown, P. J., C. D. Kummerow, and D. L. Randel, 2016: Hurricane GPROF: An optimized ocean microwave rainfall retrieval for tropical cyclones. *J. Atmos. Oceanic Technol.*, **33**, 1539–1556, doi: 10.1175/JTECH-D-15-0234.1.

Cecil, D. J. and T. Chronis, 2018: Polarization-corrected temperatures for 10-, 19-, 37-, and 89-GHz passive microwave frequencies. *J. Appl. Meteor. Climatol.*, **57**, 2249–2265, doi: 10.1175/JAMC-D-18-0022.1.

Chen, S. S., J. A. Knaff, and F. D. Marks, Jr., 2006: Effects of vertical wind shear and storm motion on tropical cyclone rainfall asymmetries deduced from TRMM. *Mon. Wea. Rev.*, **134**, 3109–3208.

Corbosiero, K. L. and J. Molinari, 2002: The effects of vertical wind shear on the distribution of convection in tropical cyclones. *Mon. Wea. Rev.*, **130**, 2110–2123, doi: https://doi.org/10.1175/1520-0493(2002)130<2110:TEOVWS>2.0.CO;2.

DeHart, J. C., R. A. Houze, Jr., and R. F. Rogers, 2014: Quadrant distribution of tropical cyclone inner-core kinematics in relation to environmental shear. *J. Atmos. Sci.*, **71**, 2713–2732, doi: 10.1175/JAS-D-13-0298.1.

DeMaria, M. and J. Kaplan, 1994: A statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic Basin. *Wea. Forecasting*, **9**, 209–220, doi: 10.1175/1520-0434(1994)009<0209:ASHIPS>2.0.CO;2.

DeMaria, M., M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan, 2005: Further improvements to the statistical hurricane intensity prediction scheme SHIPS. *Wea. Forecasting*, **20**, 531–543, doi: 10.1175/WAF862.1.

Didlake, A. C., Jr., P. Reasor, R. Rogers, and W. Lee, 2018: Dynamics of the transition from spiral rainbands to a secondary eyewall in Hurricane Earl (2010). *J. Atmos. Sci.*, **75**, 2909–2929, doi: 10.1175/JAS-D-17-0348.1.

Guimond, S. R., G. M. Heymsfield, and F. J. Turk, 2010: Multiscale observations of Hurricane Dennis (2005): The effects of hot towers on rapid intensification. *J. Atmos. Sci.*, **67**, 633–654, doi: 10.1175/2009JAS3119.1.

Hence, D. A. and R. A. Houze, Jr., 2011: Vertical structure of hurricane eyewalls as seen by the TRMM Precipitation Radar. *J. Atmos. Sci.*, **68**, 1637–1652, doi: 10.1175/2011JAS3578.1.

Hence, D. A. and R. A. Houze, Jr., 2012: Vertical structure of tropical cyclone rainbands as seen by the TRMM precipitation radar. *J. Atmos. Sci.*, **69**, 2644–2661, doi: 10.1175/JAS-D-11-0323.1.

Hendricks, E. A., M. S. Peng, B. Fu, and T. Li, 2010: Quantifying environmental control on tropical cyclone intensity change. *Mon. Wea. Rev.*, **138**, 3243–3271, doi: 10.1175/2010MWR3185.1.

Hersbach, H., et al., 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **144**, 1999–2049, doi: 10.1002/qj.3803.

Hong, Y., C. D. Kummerow, and W. S. Olson, 1999: Separation of convective and stratiform precipitation using microwave brightness temperature. *J. Appl. Meteor.*, **38**, 1195–1213.

Hou, A. Y., et al., 2014: The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, doi: 10.1175/BAMS-D-13-00164.1.

Iguchi, T., et al., 2018: GPM/DPR Level-2 Algorithm Theoretical Basis Document. NASA.

Islam, T., P. K. Srivastava, Q. Dai, M. Gupta, and W. Z. Wan Jaafar, 2015: Stratiform/convective rain delineation for trmm microwave imager. *J. Atmos. Sol.-Terr. Phys.*, **133**, 25–35, doi: 10.1016/j.jastp.2015.07.009.

Jiang, H., 2012: The relationship between tropical cyclone intensity change and the strength of inner-core convection. *Mon. Wea. Rev.*, **140**, 1164–1176, doi: 10.1175/MWR-D-11-00134.1.

Kieper, M. E. and H. Jiang, 2012: Predicting tropical cyclone rapid intensification using the 37 GHz ring pattern identified from passive microwave measurements. *Geophys. Res. Lett.*, **39**, L13 804, doi: 10.1029/2012GL052115.

Knaff, J. A., S. P. Longmore, and D. A. Molenar, 2014: An objective satellite-based tropical cyclone size climatology. *J. Climate*, **27**, 455–476, doi: 10.1175/JCLI-D-13-00096.1.

Knaff, J. A., C. R. Sampson, and G. Chirokova, 2017: A global statistical-dynamical tropical cyclone wind radii forecast scheme. *Wea. Forecasting*, **32**, 629–644, doi: 10.1175/WAF-D-16-0168.1.

Knaff, J. A., C. J. Slocum, and K. D. Musgrave, 2019: Quantification and exploration of diurnal oscillations in tropical cyclones. *Mon. Wea. Rev.*, **147**, 2105–2121, doi: 10.1175/MWR-D-18-0379.1.

Knaff, J. A., C. J. Slocum, K. D. Musgrave, C. R. Sampson, and B. R. Strahl, 2016: Using routinely available information to estimate tropical cyclone wind structure. *Mon. Wea. Rev.*, **144**, 1233–1247, doi: 10.1175/MWR-D-15-0267.1.

Knapp, K. R. and M. C. Kruk, 2010: Quantifying interagency differences in tropical cyclone best track wind speed estimates. *Mon. Wea. Rev.*, **138**, 1459–1473, doi: 10.1175/2009MWR3123.1.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.*, **15**, 809–817.

Kummerow, C. D., D. L. Randel, M. Kulie, N.-Y. Wang, R. Ferraro, S. J. Munchak, and V. Petkovic, 2015: The evolution of the Goddard profiling algorithm to a fully parametric scheme. *J. Atmos. Oceanic Technol.*, **32**, 165–176, doi: 10.1175/JTECH-D-15-0039.1.

NASA, 2021: GPM Core Observatory. Accessed 3 August 2021, `https://gpm.nasa.gov/missions/GPM/core-observatory`.

NASA Earth Data, 2021: Data processing levels. Accessed 4 November 2021, `https://earthdata.nasa.gov/collaborate/open-data-services-and-software/data-information-policy/data-levels`.

Olson, W. S., Y. Hong, C. D. Kummerow, and J. Turk, 2001: A texture-polarization method for estimating convective-stratiform precipitation area coverage from passive microwave radiometer data. *J. Appl. Meteor.*, **40**, 1577–1591.

Passive Microwave Algorithm Team Facility, 2018: Global Precipitation Measurement (GPM) mission algorithm theoretical basis document. NASA.

Pedregosa, F., et al., 2011: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830.

Petković, V., M. Orescanin, P. Kirstetter, C. Kummerow, and R. Ferraro, 2019: Enhancing PMW satellite precipitation estimation: detecting convective class. *J. Atmos. Oceanic Technol.*, **36**, 2349–2363, doi: 10.1175/JTECH-D-19-0008.1.

Petty, G., 1994: Physical retrievals of over-ocean rain rate from multichannel microwave imagery. Part I: theoretical characteristics of normalized polarization and scattering indices. *Meteor. Atmos. Phys.*, **54**, 79–99.

Razin, M. N. and M. M. Bell, 2021: The unconventional eyewall replacement cycle of Hurricane Ophelia (2008). *Mon. Wea. Rev.*, **149**, 2151–2170, doi: 10.1175/MWR-D-20-0181.1.

Razin, M. N., C. J. Slocum, J. A. Knaff, P. J. Brown, and M. M. Bell, 2021: Tropical cyclone precipitation, infrared, microwave, and environmental dataset (TC PRIMED). *Bull. Amer. Meteor. Soc.*, in peer review.

Reasor, P. D., R. Rogers, and S. Lorsolo, 2013: Environmental flow impacts on tropical cyclone structure diagnosed from airborne Doppler radar composites. *Mon. Wea. Rev.*, **141**, 2949–2969, doi: 10.1175/MWR-D-12-00334.1.

Rogers, R., 2010: Convective-scale structure and evolution during a high-resolution simulation of tropical cyclone rapid intensification. *J. Atmos. Sci.*, **67**, 44–70, doi: 10.1175/2009JAS3122.1.

Rogers, R., P. D. Reasor, and S. Lorsolo, 2013: Airborne Doppler observations of the inner-core structural differences between intensifying and steady-state tropical cyclones. *Mon. Wea. Rev.*, **141**, 2970–2991, doi: 10.1175/MWR-D-12-00357.1.

Rogers, R. F., P. D. Reasor, J. Zawislak, and L. T. Nguyen, 2020: Precipitation processes and vortex alignment during the intensification of a weak tropical cyclone in moderate vertical shear. *Mon. Wea. Rev.*, **148**, 1899–1929, doi: 10.1175/MWR-D-19-0315.1.

Rogers, R. F., P. D. Reasor, and J. A. Zhang, 2015: Multiscale structure and evolution of Hurricane Earl (2010) during rapid intensification. *Mon. Wea. Rev.*, **143**, 536–562, doi: 10.1175/MWR-D-14-00175.1.

Rogers, R. F., J. A. Zhang, J. Zawislak, H. Jiang, G. R. Alvey III, E. J. Zipser, and S. N. Stevenson, 2016: Observations of the structure and evolution of Hurricane Edouard (2014) during intensity change. Part II: kinematic structure and the distribution of deep convection. *Mon. Wea. Rev.*, **144**, 3355–3376, doi: DOI: 10.1175/MWR-D-16-0017.1.

Sampson, C. R. and A. J. Schrader, 2000: The automated tropical cyclone forecasting system (version 32). *Bull. Amer. Meteor. Soc.*, **81**, 1231–1240, doi: 10.1175/1520-0477(2000)081<1231:TATCFS>2.3.CO;2.

Schubert, W. H. and J. A. Hack, 1982: Inertial stability and tropical cyclone development. *J. Atmos. Sci.*, **39**, 1687–1697.

Shapiro, L. J. and H. E. Willoughby, 1982: The response of balanced hurricanes to local sources of heat and momentum. *J. Atmos. Sci.*, **39**, 378–394.

SHIPS, 2020: SHIPS statistical tropical cyclone intensity forecast technique development, developmental data. Accessed 3 April 2020, `https://rammb.cira.colostate.edu/research/tropical_cyclones/ships/developmental_data.asp`.

Slocum, C. J., J. P. Stow, M. N. Razin, J. A. Knaff, and A. O. Gonzalez, 2021: A new era of resolving tropical cyclones: Are renalayses there yet? *J. Appl. Meteor. Climatol.*, manuscript in preparation.

Smith, R. K., M. T. Montgomery, and V. S. Nguyen, 2009: Tropical cyclone spin-up revisited. *Quart. J. Roy. Meteor. Soc.*, **135**, 1321–1335, doi: 10.1002/qj.428.

Steiner, M., R. A. Houze, Jr., and S. E. Yuter, 1995: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. *J. Appl. Meteor.*, **34**, 1978–2007.

Stevenson, S. N., K. L. Corborsiero, and J. Molinari, 2014: The convective evolution and rapid intensification of Hurricane Earl (2010). *Mon. Wea. Rev.*, **142**, 4364–4380, doi: 10.1175/MWR-D-14-00078.1.

Susca-Lopata, G., J. Zawislak, E. J. Zipser, and R. F. Rogers, 2015: The role of observed environmental conditions and precipitation evolution in the rapid intensification of Hurricane Earl (2010). *Mon. Wea. Rev.*, **143**, 2207–2223, doi: 10.1175/MWR-D-14-00283.1.

Tao, C. and H. Jiang, 2015: Distributions of shallow to very deep precipitation-convection in rapidly intensifying tropical cyclones. *J. Climate*, **28**, 8791–8824, doi: 10.1175/JCLI-D-14-00448.1.

Tao, C., H. Jiang, and J. Zawislak, 2017: The relative importance of stratiform and convective rainfall in rapidly intensifying tropical cyclones. *Mon. Wea. Rev.*, **145**, 795–809, doi: 10.1175/MWR-D-16-0316.1.

Van Rossum, G. and F. L. Drake, 2009: *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA.

Varma, A. K. and G. Liu, 2010: On classifying rain types using satellite microwave observations. *J. Geophys. Res.*, **115**, D07 204, doi: 10.1029/2009JD012058.

Vigh, J. L. and W. H. Schubert, 2009: Rapid development of the tropical cyclone warm core. *J. Atmos. Sci.*, **66**, 3335–3350, doi: 10.1175/2009JAS3092.1.

Wadler, J. B., R. F. Rogers, and P. D. Reasor, 2018: The relationship between spatial variations in the structure of convective bursts and tropical cyclone intensification as determined by

airborne Doppler radar. *Mon. Wea. Rev.*, **146**, 761–780, doi: 10.1175/MWR-D-17-0213.1.

Zagrodnik, J. P. and H. Jiang, 2014: Rainfall, convection, and latent heating distributions in rapidly intensifying tropical cyclones. *J. Atmos. Sci.*, **71**, 2789–2809, doi: 10.1175/JAS-D-13-0314.1.