2013-03

AN EVALUATION OF NORTHERN HEMISPHERE MERGED CLOUD ANALYSES FROM THE UNITED STATES AIR FORCE CLOUD DEPICTION FORECASTING SYSTEM II

Pasillas, Chandra M.

Monterey, California. Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

AN EVALUATION OF NORTHERN HEMISPHERE MERGED CLOUD ANALYSES FROM THE UNITED STATES AIR FORCE CLOUD DEPICTION FORECASTING SYSTEM II

by

Chandra M. Pasillas

March 2013

Thesis Advisor: Tom Murphree
Co-Advisor: Tim Nobis

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** March 2013	**3. REPORT TYPE AND DATES COVERED** Master's Thesis

4. TITLE AND SUBTITLE
AN EVALUATION OF NORTHERN HEMISPHERE MERGED CLOUD ANALYSES FROM THE UNITED STATES AIR FORCE CLOUD DEPICTION FORECASTING SYSTEM II

5. FUNDING NUMBERS
RLGAJ

6. AUTHOR(S) Chandra M. Pasillas

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Reconnaissance Office (NRO) and Air Force Weather Agency (AFWA)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____N/A____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

Data from the CloudSat cloud profiling radar was used to verify the performance, or operational health, of the United States Air Force's World Wide Merged Cloud Analysis (WWMCA) system to detect clouds. WWMCA performance for 2010 over the Northern Hemisphere was analyzed by (1) cloud event category: Cloudy, Partly Cloudy, Clear; (2) geographic region: (Northern Hemisphere, 0N–50N, 0N–23.5N, 23.5N–35N, 35N-50N, 50N–90N, South China Sea, and Southwest Asia), 3) month; and (4) age of input data used by WWMCA. Overall and cloud category performance were evaluated using 11 performance metrics. Overall, WWMCA properly identified cloud categories 65% of the time, with a detection range of 40%–70% depending on the region. WWMCA performed best (worst) at low (high) latitudes. Decreases in WWMCA accuracy were noted when using input data older than 45 minutes. We confirmed that newer (older) data performed best (worst), with an improvement of nearly 20% when using all data available rather than data older than 3 hours. Annual hemispheric average Heidke skill scores were 0.52 for Cloudy, 0.47 for Clear, and 0.09 for Partly Cloudy conditions. Maximum (minimum) HSS values for all three cloud categories occurred at low (high) latitudes.

14. SUBJECT TERMS WWMCA, CLOUDSAT, CLOUD VERIFICATION, METEOROLOGY, CDFS II

15. NUMBER OF PAGES 193

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

AN EVALUATION OF NORTHERN HEMISPHERE MERGED CLOUD ANALYSES FROM THE UNITED STATES AIR FORCE CLOUD DEPICTION FORECASTING SYSTEM II

Chandra M. Pasillas
Captain, United States Air Force
B.S., United States Air Force Academy, 2004

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY

from the

NAVAL POSTGRADUATE SCHOOL
March 2013

Author: Chandra M. Pasillas

Approved by: Tom Murphree
Thesis Advisor

Tim Nobis
Thesis Co-Advisor, Air Force Weather Agency

Wendell Nuss
Chair, Department of Meteorology

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Data from the CloudSat cloud profiling radar was used to verify the performance, or operational health, of the United States Air Force's World Wide Merged Cloud Analysis (WWMCA) system to detect clouds. WWMCA performance for 2010 over the Northern Hemisphere was analyzed by (1) cloud event category: Cloudy, Partly Cloudy, Clear; (2) geographic region: (Northern Hemisphere, 0N–50N, 0N–23.5N, 23.5N–35N, 35N-50N, 50N–90N, South China Sea, and Southwest Asia), 3) month; and (4) age of input data used by WWMCA. Overall and cloud category performance were evaluated using 11 performance metrics. Overall, WWMCA properly identified cloud categories 65% of the time, with a detection range of 40%–70% depending on the region. WWMCA performed best (worst) at low (high) latitudes. Decreases in WWMCA accuracy were noted when using input data older than 45 minutes. We confirmed that newer (older) data performed best (worst), with an improvement of nearly 20% when using all data available rather than data older than 3 hours. Annual hemispheric average Heidke skill scores were 0.52 for Cloudy, 0.47 for Clear, and 0.09 for Partly Cloudy conditions. Maximum (minimum) HSS values for all three cloud categories occurred at low (high) latitudes.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AER	Atmospheric Environmental Research, Inc.
AFWA	Air Force Weather Agency
AFRL	Air Force Research Laboratory
AVDCLD	Advect Cloud Model
AVHRR	Advanced Very High Resolution Radiometer
B	bias
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CDFS	Cloud Depiction Forecast System
CPR	cloud profiling radar
CSI	critical success index
CSV	comma separated value
DCF	diagnostic cloud forecast model
DMSP	Defense Meteorological Satellite Program
DoD	Department of Defense
EO	electro-optical
ESSP	Earth Systems Science Pathfinder
F	false alarm rate
FAR	false alarm ratio
FOV	field of view
GeoProf	Geometrical Profiling
H	hit rate
HP	hit proportion
HDF	hierarchical data format
HSS	Heidke skill score
IC	intelligence community
IFOV	instantaneous field of view
IR	infrared
ISR	intelligence surveillance and reconnaissance
GDR	gridded data records
GFS	Global Forecast System
GOES	Geostationary Operational Environmental Satellite
GT	greater than
GTE	greater than or equal to
L	local time
LT	less than
LTE	less than or equal to
LWIR	longwave infrared
METEOSAT	European Space Agency's Meteorological Satellite
METAR	Meteorological Aerodrome Report
MODIS	Moderate Resolution Imaging Spectroradiometer
MSL	mean sea level
MP	miss proportion

NASA	National Aeronautics and Space Administration
NCEP	Nation Centers for Environmental Prediction
NGIS	Northrop Grumman Information Systems
NH	Northern Hemisphere
NOAA	National Oceanic Atmospheric Administration
NOGAPS	Navy Operational Global Atmospheric Prediction System
NPS	Naval Postgraduate School
NRO	National Reconnaissance Office
NWP	Numerical Weather Prediction
OCO	Orbiting Carbon Observatory
OI	optimum interpolation
OLS	operational line scanner
PARASOL	Polarization and Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar
PC	proportion correct
POFD	probability of false detection
POD	probability of detection
POES	Polar Orbiting Environmental Satellites
RRV	range resolution volume
RTNEPH	Real Time Nephanalysis
SCFM	Stochastic Cloud Forecast Model
SCS	South China Sea
SDR	sensor data records
SERCAA	Support of Environmental Requirements for Cloud Analysis and Archive
SWA	Southwest Asia
SYNOPIC	Surface Synoptic Observations
TS	threat score
UCAR	University Corporation for Atmospheric Research
USAF	United States Air Force
VT	valid time
WWMCA	Worldwide Merged Cloud Analysis

ACKNOWLEDGMENTS

I would like to thank Dr. Tim Nobis of the Air Force Weather Agency (AFWA) and Dr. Tom Murphree of the Naval Postgraduate School for their assistance, support, and dedication throughout this process. Their expertise and guidance proved extremely beneficial to the progression of this study. Without them, I would not have had the proper scientific focus to guide this study in the direction necessary to maximize the impacts of the results. I would also like to extend my gratitude to Mr. Bruce Ford of Clear Science and Mary Jordan at NPS. Without them, data matching, processing, and reduction would have been an insurmountable task for a study of this scope and size. I would also like to thank all those individuals that have assisted me in one way or another along this journey, to include Becky Selin of AFWA and Capt Morton Bartlett, as it is the sum of all of the parts that make the whole end result possible. Finally, I would like to give thanks to my husband, Gustavo Pasillas Jr., for his patience with me through this endeavor. This research project was supported by funding from the National Reconnaissance Office (NRO) and AFWA.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. IMPORTANCE OF CLOUD ANALYSES

With the constant downsizing of the Department of Defense (DoD) forces, the ever increasing threat to national security, and the constant pressure to "do more with less", it is impossible to have continuous human oversight of all of the information being received from intelligence, surveillance, and reconnaissance (ISR) systems, weather satellites, and other information sources. This is especially true for information about clouds, which are an important environmental factor in planning and executing many DoD and intelligence community (IC) operations, especially data collection by many types of ISR satellite sensors. Because of this, automated cloud detection, analysis, and forecasting processes must be developed, tested, and used in, for example, the targeting process for satellite based ISR. Improved information about clouds can help decrease the time required during the decision analysis and targeting cycle, and help increase operational efficiency and success rates. This helps ensure that a balance of weather and security factors is considered when determining the targets to be observed by a satellite.

Improved accuracy in automated cloud analyses and forecasts would lead to increased user confidence in those products for ISR satellite tasking. For example, if a cloud analysis and forecasting system with a proven level of high performance predicted unfavorable cloud conditions over a user's priority one and priority two targets, but favorable conditions over targets three, four, and five, then a satellite could be refocused on targets three through five. This would improve the chances of some successful data collection versus no data collection on the five targets because an attempt was made to collect on target one and two. Higher confidence in the cloud analyses and forecasts would allow planners to have higher confidence in their decisions to reprioritize their target list based on anticipated weather conditions. A major goal for our study was to contribute

to the improvement of cloud analyses and forecasts, and to thereby improve the planning and outcomes of DoD and IC operations.

The United States Air Force (USAF) currently uses the Cloud Depiction Forecast System (CDFS) II to combine weather satellite, surface observations, and ground characteristics to produce an hourly global cloud analysis and short term forecasts for the remote sensing intelligence community (IC). The World Wide Merged Cloud Analysis (WWMCA) is an analysis generated every thirty minutes by as part of the CDFS II process. Detailed information on the history and processes in CDFS II and WWMCA can be found in Chapter II Section B.

The goal of our study was to assess the accuracy of WWMCA in order to provide improved understanding of WWMCA and the forecasts based on WWMCA. Specifically, we conducted a detailed investigation into the accuracy of WWMCA in overall cloud detection, as well as in specific cloud categories of interest to the satellite intelligence community. Our aim was to develop objective quantitative measures of WWMCA performance that could be used to (a) assist WWMCA users in determining how to weight WWMCA and corresponding forecasts in their planning cycles; and (b) identify methods for improving WWMCA tuning and merging algorithms and the corresponding CDFS II forecasts.

B. PRIOR RESEARCH

Prior researchers have compared various cloud observation data sets to WWMCA analyses and other merged cloud analyses processes. The sources of the observational data have included surface observations, surface and air based lidar, atmospheric profilers, rawindsonds, radar, Moderate Resolution Imaging Spectroradiometer (MODIS), and CloudSat. A number of these studies have been well summarized in a Northrop Grumman Information Systems review (NGIS; 2011), including Heidemen, Ruggiero, Norquist, UCAR/AER, Gustafson and a climatology effort by the USAF 14th Weather Squadron. Table 1, at the

end of this section, summarizes the key similarities and differences between the prior studies.

Heidemen (1995) is one of the earlier studies relevant to this research project. This study reviewed data over a 10-day period in 1993 for three geographic regions of interest comparing the SERCAA analysis to the RTNEPH to determine if SERCAA should replace the RTNEPH process. Cloud coverage categories were divided into the categories (clear <20%, partly cloudy 20–80%, and cloudy >80%), and the SERCAA and RTNEPH data were compared to determined when they matched and when there was a level-2 difference (i.e., a level-2 difference occurred when one data set showed clear while the other indicated cloud). When differences occurred, human visual inspection was used to compare the analyses to the corresponding satellite imagery to see whether SERCAA or RTNEPH was more accurate. This comparison was conducted separately for day and night periods. Overall, Heidemen determined that SERCAA provided a more accurate analysis than RTNEPH, especially during daylight hours. This study provided useful information critical to developing upgrades to CDFS that were implemented in CDFS II (NGIS 2011).

Ruggiero (2000) was a validation trial of SERCAA by the Air Force Research Laboratory (AFRL). The purpose of this study was to determine if SERCAA could be used to improve the initialization of moisture fields in numerical weather prediction (NWP) modeling. SERCAA data was divided into two sets, one for data that were processed using Phase I algorithms and one for data that were processed using Phase II algorithms. Phase I algorithms were the original algorithms developed under SERCAA, while Phase II algorithms included upgrades to address shortcomings in Phase I algorithms. SERCAA Phase I and Phase II data was verified against surface based observations and rawinsondes from September 1995 in eastern Massachusetts, with comparisons being done in four cloud coverage categories: clear <0.1; scattered 0.1–0.5; broken 0.6–0.9; and overcast > 0.9. Ruggiero anticipated that data that underwent Phase II processing would perform better than Phase I, if the Phase II upgrades were

successful. Ruggiero determined that the Phase II algorithms indeed performed better than the Phase I algorithms. Phase II algorithms detected cloud when it was present 81% of the time, agreed with observed cloud fraction in 73% of the cases, was within one category in 94% of the cases, and was off by two categories in 6% of the cases (NGIS 2011). On the basis of hit rate, Ruggiero determined that using the Phase II SERCAA data improved the cloud analyses and forecasts.

Norquist (2007) verified WWMCA data for cloud layering against cloud profiling radar and portable lidar data from two studies conducted from Oct 2004 through Dec 2005 for the WWMCA cell located over Hanscom AFB, MA. In this study, data for 117 hours in which cloud cover was present was reviewed to compare WWMCA low, middle, and high cloud "observations" to radar and lidar observations in order to assess the ability of WWMCA to recognize cloud layers and levels. Norquist found that WWMCA under diagnosed high cloud and over diagnosed low and middle cloud. Overall, Norquist determined that WWMCA did a good job of cloud detection and its performance was reasonable for low and middle level clouds (NGIS 2011).

Horsman (2007) conducted a study using data from ten Air Force bases across the continental United States for 16 days in 2007. Hourly manual observations were compared to the ten corresponding WWMCA cell values. Horsman determined that the WWMCA had a poor performance for point locations, with an overall verification rate of 27% and a miss rate of 32% (Horseman 2007).

University Corporation for Atmospheric Research (UCAR), conducted a comparative study of WWMCA to CloudSat in 2008 and the Atmospheric Environmental Research, Inc. (AER), followed up these findings with a similar study in 2010. These studies were done with the assistance of the U.S. Air Force 16th Weather Squadron. No formal reports were released but the studies were summarized in a verification and validation paper by Northrop Grumman Information Systems in 2011(NGIS 2011). The first study covered the Northern

Hemisphere from 1 April–29 June 2008 using only WWMCA data less than three hours old. The 2010 follow up study included the original analyzed period, a secondary three month period from 28 March–31 May 2010, and expanded to include both hemispheres. It studied performance using all WWMCA data regardless of age and using only data less than three hours old. CloudSat and WWMCA data sets are at different resolutions, so the two data points had to be matched to similar spatial and temporal periods for comparison. All CloudSat data sixty minutes prior to the hour was consolidated to be included for comparison to the top of the hour WWMCA analysis. It was then converted into WWMCA cell coordinates and binned accordingly. If more than six CloudSat pixels were included in a WWMCA cell, then the CloudSat cloud percentage and cloud category was calculated for direct comparison to the WWMCA values. More details on this process will be discussed in Chapter II. UCAR and AER determined that WWMCA had a tendency to under analyze cloud. Results from using all data available showed that WWMCA detected cloud when CloudSat said there was cloud more than 80% of the time and agreed on the cloud/no cloud amounts 75% of the time (NGIS 2011).

Bartlett (2009) conducted qualitative studies to compare WWMCA to data from MODIS and from the Navy Operational Global Atmospheric Prediction System (NOGAPS). While not a quantitative study, Bartlett determined that quantitative verification of cloud distributions could be achieved (Bartlett 2009).

Gustafson (2011) compared MODIS derived cloud mask data to WWMCA data on a global scale for June and September of 2010. The MODIS data was converted to “yes/no” cloud for each 1 km box and then matched to WWMCA cells. A percent cloudy calculation was done based on the number of MODIS data points in the WWMCA cell. Unlike the UCAR and AER studies, the number of MODIS points available in a WWMCA cell did not appear to be a consideration in the Gustafson study. Cloud fractions for both the MODIS cloud masks and WWMCA were calculated and classified as clear (<20%), partly-cloud (20–80%), or cloudy (>80%). Gustafson also noted when there was a greater than 20%

difference in the cloud fractions between MODIS and WWMCA. Gustafson determined that the cloud fractions of WWMCA and MODIS were in agreement 65% of the time, and that MODIS performed better over the polar regions, while WWMCA performed better over bright backgrounds, such as deserts or sun glint. Gustafson determined that, while a useful comparison, the MODIS cloud mask is less than ideal as an independent source of "truth" for analysis comparison and recommended a similar study be conducted using CloudSat as "truth" (NGIS 2011).

Stubblefield (2011) worked to extract value from ensembles for cloud free forecasting using forecasts created using WWMCA combined with output from the National Centers for Environmental Prediction (NCEP) global weather ensemble. Data over three climatologically different regions from Feb 2010–Jan 2011 was compiled and analyzed to determine Heidke skill score, true skill score (TSS), hit ratio, correct reject ratio, odds ratio, and relative operating characteristic (ROC). The focus was on determining if skill could be increased by using ensemble forecasting for clouds; however, some valuable information on WWMCA performance was discovered as well. Stubblefield found that the tendency for WWMCA to create bi-modal distribution of 100 or 0 percent clear or cloudy conditions was problematic for both forecasting and verification. Specific to our study, Stubblefield determined that appreciable skill and value does exist in WWMCA cloud free forecasts and that the skill varies with cloud type and frequency. Stubblefield also determined that the advection scheme of ADVCLD gives a poor representation of cloud cover processes and evolution and should be re-evaluated (Stubblefield 2011).

Cleary (2012) conducted a study using the CloudSat 2B-GeoProf data mask to verify the accuracy of WWMCA for southwest Asia and western Russia for January, April, July, and October 2010. Cleary used methods similar to the UCAR (2008) and AER (2010) studies for spatial and temporal matching of CloudSat to WWMCA. Contingency table metrics were then calculated and used to compare WWMCA performance in three cloud cover categories: clear (<20%

cloud), probably cloud (20–80% cloud), and cloud (> 80% cloud) for day and night performances. He used the CloudSat data mask values of 20 and 30 as cloud occurrence thresholds. Cleary found that WWMCA performed better in persistent cloud conditions than variable cloud conditions, and that WWMCA performed poorly in partly cloudy conditions. HSS values were higher for mid-latitude analyses than high latitude analyses. Cleary found that overall WWMCA performance was lower than that discovered in prior studies. Additionally, Cleary concluded that WWMCA performance was relatively insensitive to the 20 and 30 CloudSat cloud masking thresholds used in the verification (Cleary 2012a).

Table 1. Summary and comparison of prior research to include key studies and their data sets, verification metrics used, results, and conclusions.

Study	Period	Region	Data Sets Compared	Verifying Data Period	Testing Method	Verification Metrics	Results and Conclusions
Heidemen (1995)	May 93 (10 days)	Three regions: Japan, Central America, Himalayas	SERCCA vs RTNEPH	N/A	Compared values for three cloud categories (LT 20%, 20–80%, GT 80%)	Hit rate Level-2 difference	SERCAA outperformed RTNEPH
Ruggiero (2000)	Sept 95 (48 cases)	One Region: Eastern Massachusetts	SERCAA vs Surface Observations and Rawinsonde Data	N/A	Compared values for four cloud categories (clear, sct, bkn, ovc)	Cloud detection rate	SERCAA correctly detected cloud 81%, cloud fraction in agreement 73%, cloud fraction was within one category 94%, cloud fraction significantly disagreed 6%
Norquist (2007)	Oct 04 – Feb 05 (10 days) and Feb – Dec 05 (26 days): total 117hrs	One WWMCA cell (~24km x 24km)	WWMCA vs LIDAR	N/A	Compared values for cloud height detection	Cloud detection rate	WWMCA: detected clouds 78%, under diagnosed high cloud, over diagnosed low and middle cloud
Horseman (2007)	16 days	10 U.S. military bases	WWMCA vs Surface Observation	N/A	Compared values for YES/NO cloud (percent not specified)	Verification Miss rate	WWMCA: verification rate 27%, miss rate 32%
UCAR (2008)/AER (2010)	1 Apr – 29 Jun 08 28 Mar – 31 May 10	2008 – Northern Hemisphere 2010 - Global	WWMCA vs CloudSat 2B GeoProf	CloudSat matching to WWMCA VT- 60 min	Compared values for YES/NO cloud (hit = greater than 1% cloud), done with all WWMCA data and only WWMCA data less than 3 hrs old	Hit rate	WWMCA: detected cloud more than 80%, agreed with CloudSat 75%
Bartlett (2009)	April – June 08	Two regions: Horn of Africa and Sub-Saharan Africa	WWMCA vs MODIS and NOGAPS	Varied, typically +/- 30 min between WWMCA and MODIS	Comparison of total cloud fraction (0–20%, 20–40%,40–60%,60–80%,80–100%), and cloud layering using Euclidean and Kulback-Leibler methods	Visual quantitative and qualitative inspection	WWMCA did well at depicting large areas of distinct cloud cover at meso-beta and higher scales. Quantitative evaluation of clouds can be achieved
Gustafson (2011)	Sept and June 2010	Global	WWMCA vs MODIS	MODIS matching to WWMCA VT – 60min	Compared values for three cloud categories (Clear < 20%, Partly Cloudy 20–80%, Cloud > 80%) and greater tha n20% differences	Proportion correct Greater than 20% difference between MODIS and WWMCA	65% of the time the sky was either clear or cloudy. WWMCA and MODIS: Clear match 25%, Partly Cloudy match 6%, and Cloudy match 34%. MODIS had 20% or more cloud than WWMCA 28%, WWMCA had 20% or more cloud than MODIS 7%. WWMCA outperformed MODIS over desert regions and sun glint. MODIS outperformed WWMCA in polar regions. MODIS not 100% accurate so may not be the best "truth" for verification. Locations were MODIS and WWMCA disagree can help optimally focus tuning efforts
Stubblefield (2011)	Feb 2010 –Jan 2011	Three regions: Saudi Arabia and Iran, China, northern South America	Ensemble forecasting with WWMCA and National Centers for Environmental Predictions global weather ensemble	N/A	30% cloud/no cloud threshold	Heidke Skill Score (HSS), True skill score (TSS), relative operating characteristic (ROC), hit ratio, correct reject ratio, odds ratio	Bi-modal distribution discourages use of varied cloud fractions. Advection scheme of ADVCLD is poor representation of cloud cover processes and evolution. Utility theory should be employed operationally. Appreciable skill and value exists in cloud free forecasts and skill varies with cloud type and frequency.
Cleary (2012)	Jan, Apr, Jul and Oct 2010	Two regions: Russia (WWMCA box 26) and SWA (WWMCA box 22)	WWMCA vs CloudSat 2B GeoProf	CloudSat matching to WWMCA VT+/- 30min	Compared values for three cloud categories (LT 20%, 20–80%, GTE 80%)	Probability of detection (POD), proportion correct (PC), false alarm (FA), threat score (TS), bias, HSS	WWMCA performed better in persistent cloud conditions than in variable cloud conditions. WWMCA performance was lower than in prior studies. WWMCA performance was insensitive to CloudSat cloud mask thresholds of 20 and 30. HSS higher for mid-latitude than high latitude. WWMCA performance poor for partly cloudy conditions.

C. MOTIVATION AND SCOPE OF RESEARCH

The prior studies were critical for improving automated cloud analyses and forecast systems, but additional research is still needed to assess these systems and improve their performance. Most of the prior studies only accounted for a small localized region and/or short period of time, and thus may not be representative of performance in other regions and periods, or of overall global performance. Most of the prior studies used a very limited set of performance metrics, and thus provided only a limit performance perspective. The Ruggiero study only accounted for hit rates and cloud category differences, and did not present a full picture of the overall performance. For example, it did not explicitly provide information on false alarms ratios or rates, threat scores, and other conventional verification metrics. The Norquist (2007) study addressed the ability of WWMCA to recognize clouds at various levels versus overall cloud detection, but only considered hit rates in its verification. Horsman (2006) used both hit rate and miss rate; however, he only accounted for a single WWMCA cell for 16 days, and used a single point observation to represent an entire 24 km x 24 km region. The UCAR (2008) and AER (2010) studies expanded the areas and periods of interest to include a more global perspective and multiple months of data, which had not been done in prior studies. Unfortunately, these studies only looked at “yes/no” cloud hit rates and averaged CloudSat data up to an hour old to determine the top of the hour “truth” for comparison to WWMCA. Cleary (2012) reduced the time matching for CloudSat “truth” from the 60 min window to a 30 minute window; however, the method used for both his study and the UCAR and AER studies for spatial matching made a large assumption that six CloudSat pixels was an adequate amount of verifying data to use in assessing a given WWMCA cell. None of the prior studies accounted for latitudinal variation of the WWMCA cell size.

We addressed these shortcomings in our study by: (1) increasing the study period and study region to cover a full year and all of the northern hemisphere; (2) accounting for latitudinal variations in WWMCA cell size; (3)

setting a minimum for CloudSat coverage within a WWMCA cell of 6%; (4) using a wide range of performance metrics; (5) limiting the window of CloudSat data use to determine the "truth"; and (5) addressing the age, or time latency, of the input data used to generate WWMCA. Specifically, WWMCA performance for 2010 over the Northern Hemisphere was analyzed by: (1) cloud event category: cloudy (80–100% cloud cover), partly cloudy (20–79% cloud cover), clear (0–19% cloud cover); (2) geographic region (Northern Hemisphere, below 50 degrees north latitude, tropics (0–23.5N), subtropics (23.5–35N), mid-latitude (35-50N), high-latitude (50–90N), South China Sea, and Southwest Asia; (3) month (January–December); and (4) pixel age of WWMCA data. WWMCA overall and cloud category performance were evaluated using contingency table and other metrics, including probability of detection, probability of false detection, proportion correct, threat score, false alarm ratio, Heidke skill score, bias, hit proportion, miss proportion, and level-1 and level-2 differences. Our intention in conducting a more comprehensive study than prior studies was to develop a more robust and detailed assessment of WWMCA. Table 2 summarizes our study for comparison to prior studies (see Table 1).

Table 2. Summary of current research.

Study	Period	Region	Data Sets Compared	Verifying Data Period	Testing Method	Verification Metrics	Results and Conclusions
Pasillas (2013)	Jan – Dec 2010	Eight regions: Northern Hemisphere 0-90N, Tropics 0-23.5N, Subtropics 23.5-35N Midlatitude 35-50N. High latitude 50-90N, 0-50N, Southeast Asia WWMCA Box 22, South China Sea WWMCA Box 12	WWMCA vs CloudSat 2B GeoProf	CloudSat matching to WWMCA VT – 15min	Compared values for three cloud categories and level 1 and level 2 differences as well as latency.	Hit proportion, miss proportion, level-1 difference, level-2 difference, PC, POD, TS, FAR, POFD, bias, HSS	See Chapter III

The variety of methods seen in prior research demonstrates that there are many approaches to assessing the performance of cloud analysis and forecasting products such as WWMCA. It is important to note that there is not one generally agreed upon approach for verifying such products and for

determining what independent data sources should be used for verification purposes. The large variation in the studies also shows the complexity of interpreting performance results from similar studies for the same forecasting tools. Additionally, WWMCA is a constantly evolving product with regular upgrades, including a major upgrade in April 2009. WWMCA verification results for prior periods may not be applicable to newer updated versions of WWMCA. This is why it is difficult to do a direct comparison of results between similar studies on WWMCA performance. Ingenuity and originality are appreciated in the scientific community, but the meteorological research community would benefit greatly from some form of standard operating procedures for analysis and verification of WWMCA performance.

D. ORGANIZATION

Chapter I includes an overview of the study, summary of prior research, motivations, and scope of research. Chapter II provides a history and explanation of cloud forecasting systems and processes, information on the system being used for verification, data for analysis, data reduction and assimilation processes, and methods used to verify WWMCA performance against CloudSat. Chapter III discusses the results of the WWMCA verification study addressing both latency and overall performance. Chapter IV provides a summary of findings, conclusions, and recommendations for future research.

THIS PAGE INTENTIONALLY LEFT BLANK

II. DATA AND METHODS

A. OVERVIEW

This chapter provides background information on CDFS II, WWMCA, and CloudSat. Additionally, it discusses the data sets used in this study and outlines the methods used in data processing, assimilation, and analysis. Specifically, this chapter outlines the steps to temporally and spatially match the WWMCA and CloudSat data sets, explains the process and purposes for preliminary analyses done after data matching, and provides details on the specific methods and metrics used in the focus analyses of this study.

B. AUTOMATED CLOUD ANALYSIS

1. Cloud Depiction and Forecasting System II

Prior to CDFS II, the original CDFS used Real-Time Nephanalysis (RTNEPH) for its data processing. RTNEPH was greatly limited in the data that it could ingest and combined conventional surface observations with reduced resolution single channel infrared or visual channels from the Defense Meteorological Satellite Program (DMSP) operational line scanner (OLS) or National Oceanic Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) data to create a merged analysis of clouds. The RTNEPH model produced four floating layers of clouds and provided total and layered cloud amounts, cloud layer tops, cloud layer bases and cloud type at 48 km resolution. RTNEPH was dependent on polar orbiting satellites which greatly reduced the frequency at which it could produce reliable and timely products (Isaacs 1994).

The Support of Environmental Requirements of Cloud Analysis and Archive (SERCAA) was an initiative of the Air Force research community to improve upon the RTNEPH process and provide the next generation nephanalysis model for CDFS II. It also was used to create a new global cloud algorithm for use in determining radiative and hydrological effects of clouds on

climate and global change (Isaacs 1994). Its main goal was to improve the automated nephanalysis capabilities for multi-platform sensors. While RTNEPH was tuned to accept only one channel of infrared or visual channels, SERCAA made use of full resolution data and multiple visible and infrared channels from the various satellites. Furthermore, SERCAA allowed for three separate nephanalyses to be produced at the sensor resolution before the integration of satellites into a single analysis. The final analysis included up to four floating layers of clouds and provided total and layered cloud amounts and cloud type done at a 24 km resolution versus 48 km (Isaacs 1994). Having algorithms tuned for specific sensors capabilities improved the ability to detect, and more accurately layer, clouds over the original RTNEPH processes.

In 1998, CDFS II was released as an initiative to increase the amounts and types of satellite information ingested into cloud analysis processes in order to improve cloud detection, analyses, and forecasts. This initiative allowed for the inclusion of rapidly updating geostationary satellites to improve the temporal and spatial resolution for automated cloud analysis. Similar satellite types and systems are grouped into families and their data merged into a single gridded data record (GDR) for that sensor family and time. There are four primary sensor families and associated GDRs: Defense Meteorological Satellite Program (DMSP), Geostationary (GEO's) which contains the five geostationary platforms, Television Infrared Observation Satellite Program (TIROS), which contain the NOAA polar orbiters and the European METOP polar orbiter (all of these sources fly the same instrument) and a GDR containing the two NASA MODIS satellites. For this study there were only three GDR families (DMSP, GEOs, and TIROS) as the addition of the MODIS GDR did not occur until after 2010. During the retrieval of satellite information, SERCAA algorithms are applied to the data from the individual satellites, and cloud layers and types are determined for each input sensor and mapped to the GDR (Isaacs 1994). This is an ongoing process that creates a new GDR every time new data becomes available any sensor within that family. The typical refresh rate for a geostationary satellite is every 15–30

minutes and every 75–90 minutes for a polar orbiter. Figure 1 and Figure 2 show the data coverage provided by geostationary and polar orbiting satellites for the CDFS II in 2012. The MODIS satellite was not available in 2010. There is a gap in polar coverage from 2200–0100 local time (L) and 1000L–1300L as well as from 0300L–0400L and 1500L–1600L. This time amounts to approximately 66% of the hours not covered by a satellite in orbit.

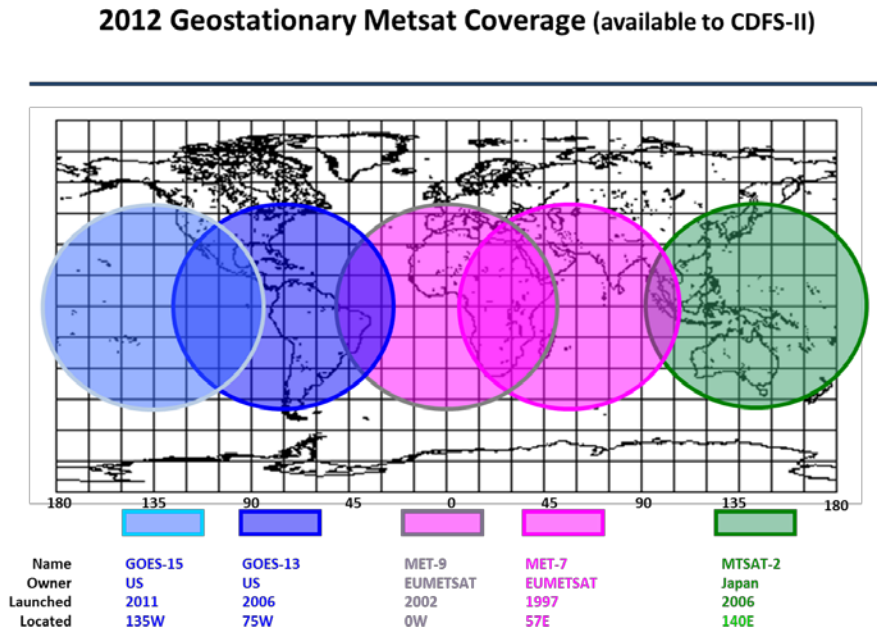


Figure 1. Geostationary Metsat coverage available to the CDFS II for 2012. Coverage extends to 50 degrees along satellite centerline and is overlapping (From AFWA 2011).

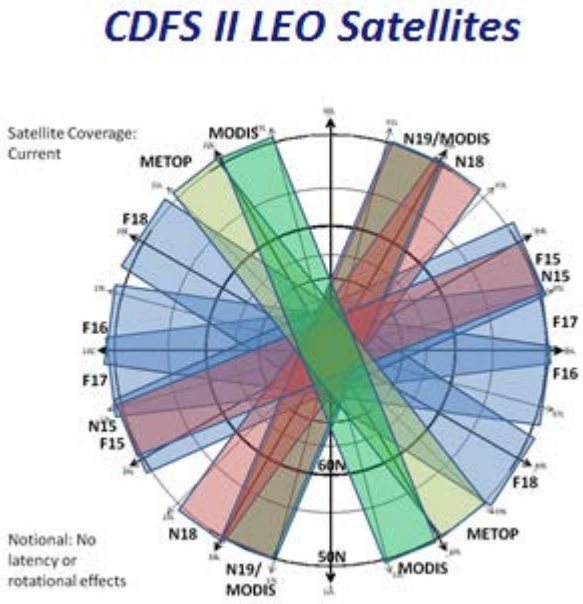


Figure 2. Polar orbiting satellite coverage available for CDFS II in 2012. The N15 and F15 satellites are at the end of their life cycle and, if not replaced, will cause more gaps, decreasing polar coverage to 66% of the day as in 2010 (From AFWA 2011).

The compiled analysis created by CDFS II is called the World Wide Merged Cloud Analysis (WWMCA). There are four main levels of processing in CDFS II to create the WWMCA. Level 1 is ingest-processing that includes unpacking of the telemetry stream, sensor data calibration, and earth location. Level 2 is cloud-detection, cloud optical property retrieval, and parallax correction performed on a pixel-by-pixel basis using sensor-specific algorithms to analyze data transmitted from each satellite. Level 3 is cloud-layering and typing to provide a vertical stratification of the cloud-filled pixels detected in Level 2. Level 3 output is remapped to the standard AFWA polar-stereographic grid projection at a resolution of 24 km (true at 60 degrees latitude). Level 1, 2, and 3 are event driven processes triggered by the receipt of new data from any of the satellite sources. Level 4 is integration, or merge, processing wherein the most recent analyzed products from each satellite and all available surface observer reports are combined to produce the World Wide Merged Cloud Analysis. (HQ/AFWA 2012)Levels 1–3 are constantly updating, while level 4 is a time driven process that occurs at the top and bottom of the hour. The bottom of the hour analysis is

distributed without modifications while the top of the hour analysis reviewed by a weather forecaster in the loop (FITL). The WWMCA that goes through forecaster manipulation is labeled as a top of the hour (e.g., 12Z) analysis of global cloud coverage but contains little to no data for that precise valid time. WWMCA is actually comprised of data that varies strongly in age from near zero minutes to greater than four hours old, due to the refresh rate and availability of data (T. Nobis 2013, personal communication). If necessary, the FITL will make slight modifications to cloud cover in focus regions based on visual comparisons to satellite imagery before the top of the hour output is used to initiate various model outputs. Model outputs may then again be adjusted by a FITL and released for use in the weather intelligence cycle. This FITL process is repeatedly every 60 minutes, 365 days a year (T. Nobis 2012, personal communication). Figure 3 depicts how CDFS II combines surface and upper air observations, specialized global analyses of surface temperatures and snow depths, and the data from the four GDRs created from each meteorological satellite family to create a merged cloud analysis. Levels 1–3 occur during the satellite tuning portion of this diagram. A more detailed summary of all of the CDFS processing levels from Cleary (2012a) can be found in Appendix A.

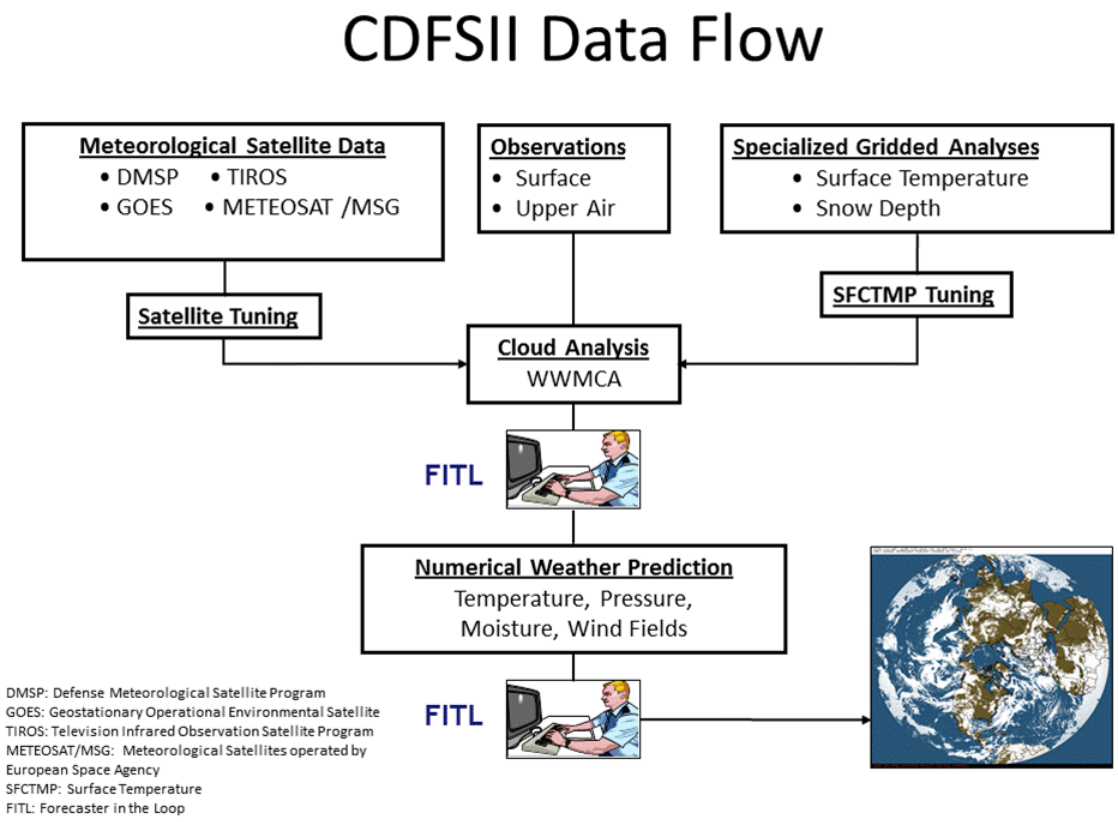


Figure 3. Flow chart for data in the CDFS II process. Satellite data, observations, and modeled surface temperature and snow depth are adjusted and merged to create the WWMCA cloud analysis. The analysis is then combined with numerical weather predictions to output cloud forecasts (After AFWA 2011).

All steps in the CDFS II process are critical; however, it is important to highlight the level 4 cloud analysis integration step in which the CDFS II merges the GDRs from each satellite sensor into a single analysis. This occurs during the cloud analysis step in Figure 3 using the integration processes outlined in Figure 4. This is the step which results from our study may help to improve. Further information on this step is provided by HQ AFWA (2012), in which the following is stated about the satellite merging process:

> Integration of total cloud amount precedes integration of layer quantities since the estimates of total cloud fraction are believed to be more reliable than any individual layer fraction (due to small

sample sizes and the potential for height assignment errors). Processing occurs independently for each grid cell. First, a series of rules is applied to determine if any one of the input analyses is superior to the other two. If none of the input grids have been updated since the time of the previous Worldwide Merged Analysis, then the previous analysis is persisted. If new analyses are available, a check is made to determine if more than one are timely. If only one timely analysis is available, the merged total cloud fraction is set to the value of this analysis. If more than one analysis satisfies timeliness requirements, these analyses are examined to determine if they are all either completely cloud-filled or completely cloud-free. If so, total cloud fraction is set to either 100 or 0 percent, respectively. If multiple timely analyses exist that are neither all completely clear nor completely cloud-filled, then an estimated error of each sensor analysis is used to determine if the most recent analysis also has the lowest estimated error. Only when all these conditions fail is an optimum interpolation (OI) algorithm used to obtain a blended estimate of total cloud fraction from multiple input analyses. Averaging weights for the OI are based on estimated analysis errors computed for each available sensor analysis. (HQ AFWA 2012)

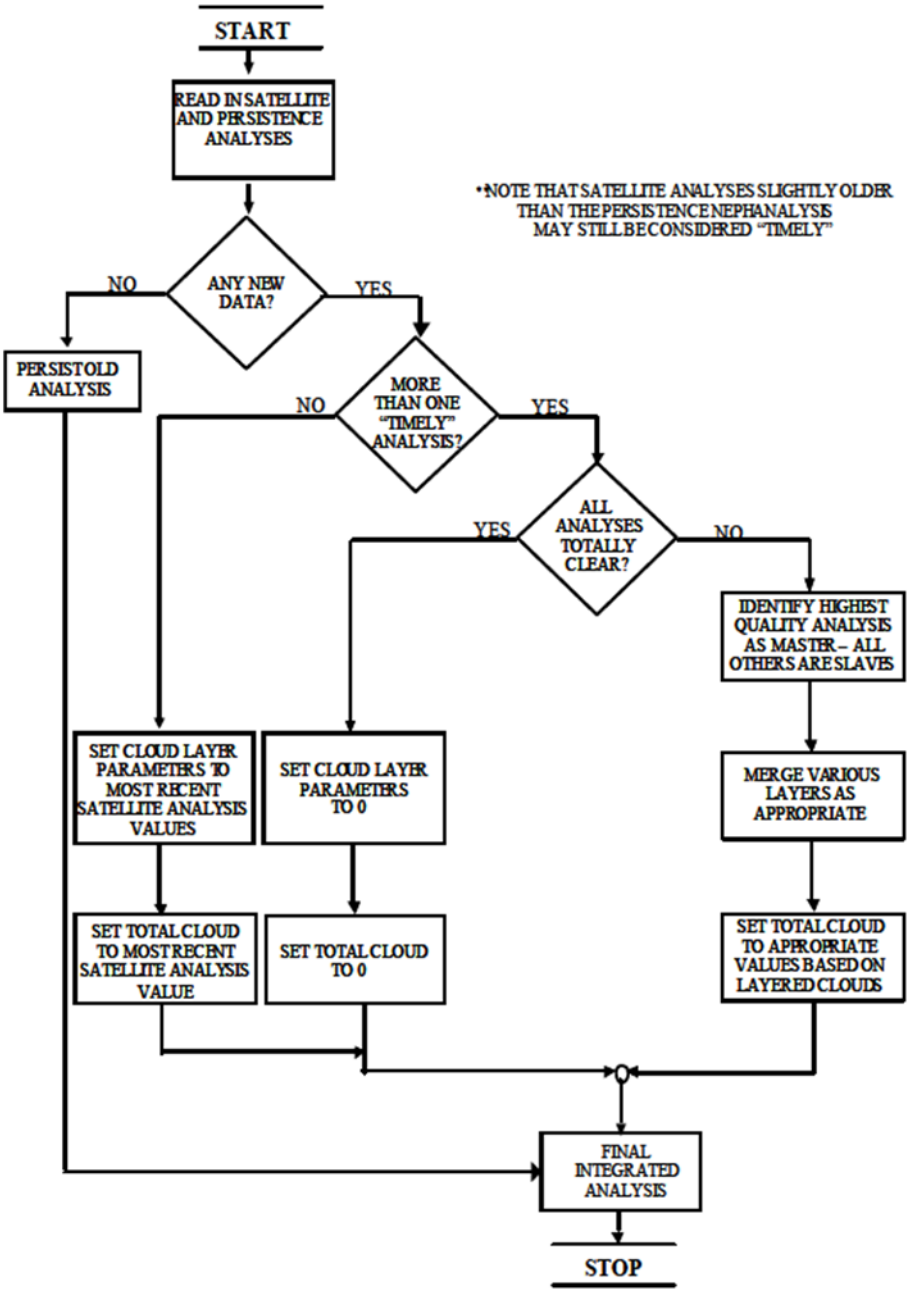


Figure 4. Cloud analysis integration functional flow (From HQ AFWA 2012).

2. World Wide Merged Cloud Analysis

WWMCA analyses are done by overlaying a regular rectangular grid on to a polar stereographic projection as seen in Figure 5. This mesh grid creates 64 numbered boxes for the Northern and Southern Hemispheres, referred to in our

study as the WWMCA boxes. These numbered boxes represent a "whole" mesh grid and the area is further reduced to a 1/16th mesh with a 1024 x 1024 grid with I,J values that represent the latitudes and longitudes of the 1/16th mesh. We refer to the cells defined by the 1024 x 1024 grid as the WWMCA cells. These smaller cells are 1/16th of the whole mesh, which translates to a cell size that is 24 km x 24 km at 60 degrees latitude. This is the resolution at which cloud classification is done for WWMCA. Using this projection and meshing, WWMCA cell size is a function of latitude, with a minimum box size of 12.5 km at the equator and at a maximum of 25 km at the poles. This resolution is an improvement over prior cloud merging processes; however, meteorological satellites often have higher resolutions, ranging from 1 to 15 km, higher than the current average WWMCA resolution (HQ AFWA 2012).

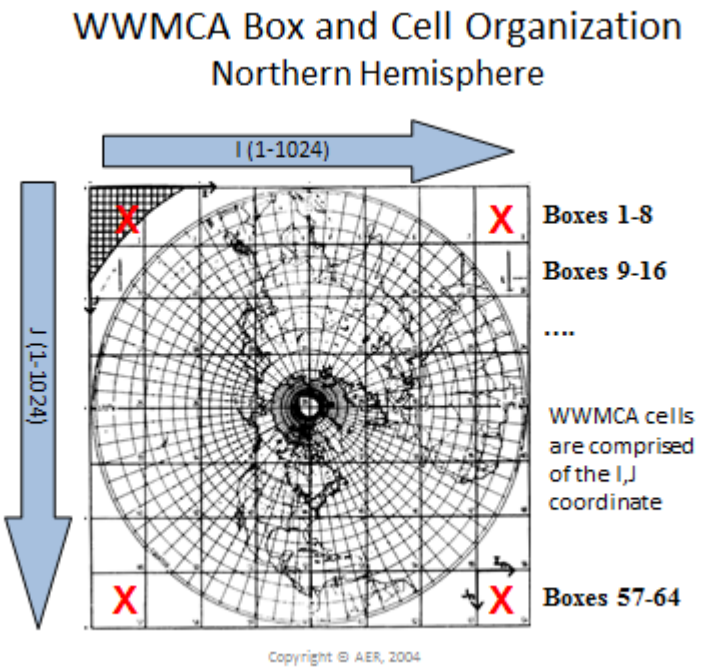


Figure 5. WWMCA box and cell orientation on a 1/16th mesh grid. The WWMCA whole mesh is represented by the boxes numbered 1-64, while the WWMCA cells are sub-regions defined by the 1/16th mesh (After Cleary 2012b).

Due to its current merging algorithm, the tendency for CDFS II is to categorize clouds in a WWMCA cell as either 100% clear or 100% cloud during the creation of GDRs and the merging process, when in actuality there may be a combination of both clear and cloud conditions that are ignored by WWMCA. The OI algorithm is rarely used to obtain a blended estimate of total cloud fraction for a WWMCA cell, even though the input satellite is at a higher spatial resolution and may indicate mixed cloud. This may lead to problems in identifying smaller features and areas of cloudiness within larger patches of clear skies, or small clearings in larger cloud shields. Figure 6 shows the distribution of WWMCA analyses of cloud cover for the Northern Hemisphere for 00Z from 12 – 21 February 2012. Notice the predominance of clear and cloud conditions, with over 70% of the WWMCA analyses indicating no cloud or cloud conditions, and less than 30% indicating partly cloudy conditions. This figure is representative of the tendency or bias in the WWMCA analyses toward clear or cloud noted in other studies (e.g., HQ AFWA 2012; Cleary 2012).

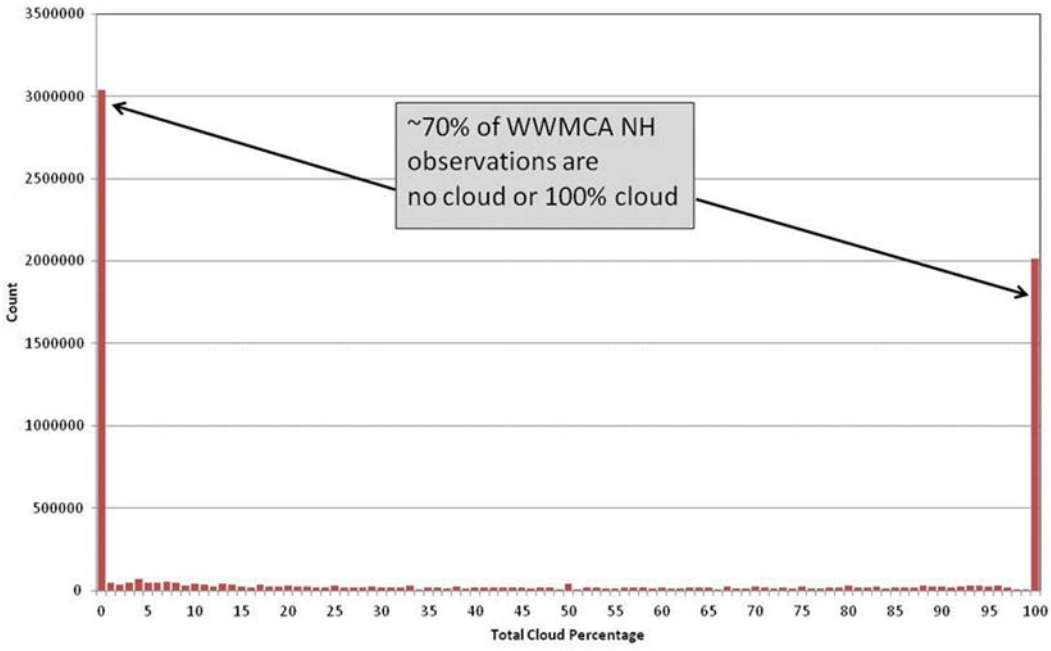


Figure 6. The distribution of WWMCA analyses for 00Z, Northern Hemisphere from 12–21 Feb 2012 (T. Nobis, 2012, personal communication).

C. CLOUDSAT

1. Description of CloudSat

CloudSat is a National Aeronautics and Space Administration (NASA) Earth Sciences Systems Pathfinder (ESSP) mission that was launched in April 2006. CloudSat's mission is to measure the vertical structure of clouds from space, and observe cloud and precipitation. The primary CloudSat instrument is a W-band (94-GHz), nadir-pointing, Cloud Profiling Radar (CPR) (NASA 2008). The original CloudSat program was funded to operate for only 22 months, but received an extension of mission operations to September 2011. As of February 2013, CloudSat is still operating, but at a reduced capability, and is now limited to day use only. Ground operations and satellite communications are performed by the USAF at Kirkland Air Force Base, Albuquerque, New Mexico, and the data is downloaded and processed at the CloudSat Data Processing Center (CDPC) at Colorado State University (CSU) (Stephens 2008). The four key mission objectives for CloudSat are as follows:

> (1) quantitatively evaluate the representation of clouds and cloud processes in global atmospheric circulation models, (2) quantitatively evaluate the relationship between the vertical profiles of cloud liquid water and ice and the radiative heating of the atmosphere and surface, (3) evaluate cloud properties retrieved from other satellite systems, in particular those of Aqua, and (4) contribute to improving our understanding of the indirect effect of aerosols on clouds by investigating the effect of aerosols on cloud and precipitation formation. (Stephens 2008)

CloudSat is just one satellite in a constellation of satellites commonly referred to as the "A-Train". The constellation flies in a Sun-synchronous orbit with a mean equatorial altitude of 705–730 km and an inclination of 98.2°. This orbit is fixed, so that there are no changes in the orbital elements over long time periods. The satellites cross the equator at approximately 1330L and 0130L every day and the revisit interval to the exact same location is 16 days. This means that CloudSat repeats its ground track every 16 days, or 233 revolutions (NASA 2011). The A-Train consists of six satellites: CloudSat, Cloud-Aerosol

Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite, Aura, Polarization and Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar (PARASOL), Aqua, and Orbiting Carbon Observatory (OCO). The MODIS sensor is flow on the Aqua. The relationship of these satellites can be seen in Figure 7. While each satellite itself provides critical information to the meteorological community, it is the synergistic effect of the satellites working together that is unparalleled by any other sensing system:

> By combining the components, scientists are able to gain a better understanding of important parameters related to climate change. The A-Train formation will allow for synergistic measurements where data from several different satellites can be used together to obtain comprehensive information about various key atmospheric components or processes. Combining the information from several sources gives a more complete answer to many questions than would be possible from any single satellite taken by itself. (NASA 2003)

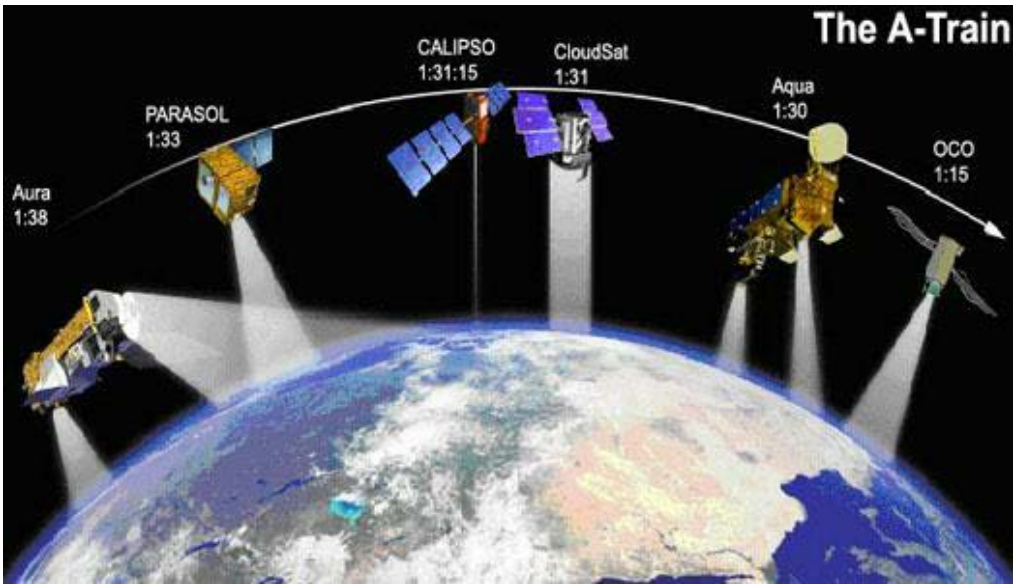


Figure 7. A depiction of the satellites that make up the "A-Train" constellation. The satellite name and equator crossing is shown for each satellite in the constellation. Note that the gap between the coverage of the first and last satellite in the "A-Train" is less than 30 minutes (and less than two minutes between Aqua, CloudSat, and CALIPSO), which allows for synergy within the constellation (From NASA 2003).

Data from a single orbit of the CloudSat is referred to as a CloudSat granule. Each granule represents a surface path that is 40,786 km in length and contains approximately 37,088 profiles. Each profile represents a vertical sounding from the satellite through the atmosphere to the surface. The horizontal surface area represented by each profile is the CloudSat pixel area or instantaneous field of view (IFOV). From an altitude of 705 km, the IFOV, at mean sea level, is 1.7 km along and 1.3 km across track (Mace et al. 2007). A dissection of the CloudSat granule can be seen in Figure 8.

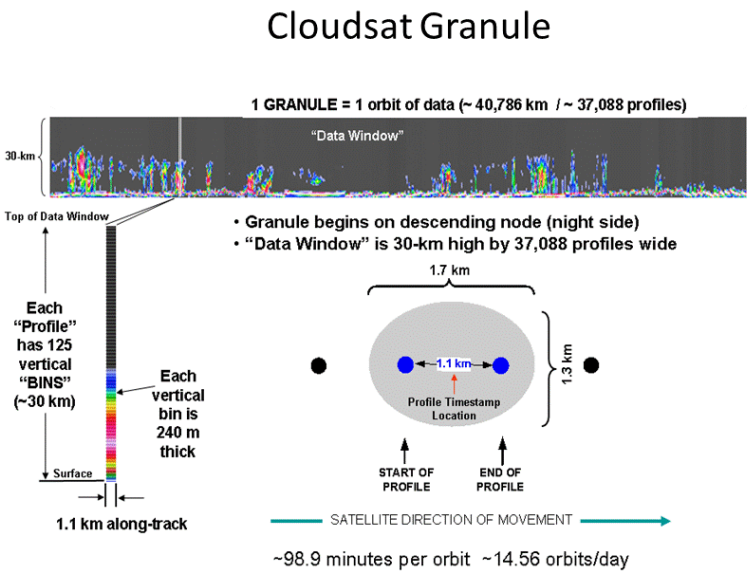


Figure 8. Dissection of a single CloudSat Granule, including the IFOV for the CloudSat cloud profiling radar (CPR). (From Cleary 2012b).

2. CPR and the 2B Cloud Geometrical Profiling (GeoProf) Product

The cloud profiling radar (CPR) is the sensor on CloudSat that is used to detect microwave radiation from clouds, and thus infer clouds and precipitation in the atmosphere. The CPR has provided nearly continuous, global time series of vertical cloud structure and properties at a vertical resolution of 485 m since 2 June 2006. The CPR emits a 3.3 microsecond pulse resulting in a vertical

resolution of 485 m. The back scattered signal is oversampled to produce a range gate spacing of 240 m (Stephens et al. 2008). This spacing leads to 124 levels in the vertical profile as see in Figure 9.

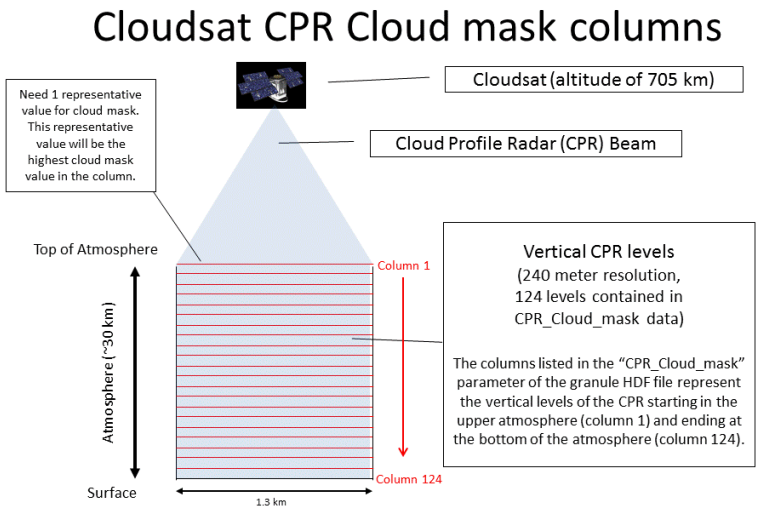


Figure 9. Description of the vertical layer resolution of the CloudSat cloud profiler (From Cleary 2012b).

There are several products derived from measurements taken by the CPR. The 2B-GeoProf product provides information about the presence or absence of clouds by identifying the levels in the vertical column sampled by CloudSat that contain significant radar echo from hydrometeors and providing an estimate of the radar reflectivity factor for each of these volumes. The cloud mask data portion of the 2B-GeoProf provides values that indicate the likelihood of cloud detection at each vertical profile level over a CloudSat pixel. The cloud mask data is stored in the 2B-GeoProf data product and contains a value between 0 and 40 for each range bin, with values greater than 5 indicating the likelihood of hydrometeors (Stephens et al. 2008). Larger values indicate a higher likelihood of hydrometeors, and hence clouds, and a lower likelihood of false detections. The measurable cloud mask values and product cautions are

described in Figure 10. For this study, we used the maximum detected value and level of occurrence to determine if there was any detection of cloud in the vertical column over each CloudSat pixel.

Discussion of Product & Product Quality

The significant echo mask is stored under the variable name "CPR_Cloud_Mask", and contains a value between 0 and 40 for each range bin with values greater than 5 indicating the location of likely hydrometeors. Increasing values of the "CPR_Cloud_Mask" variable indicate a reduced probability of a false detection, as summarized in Table 1.

Mask Value	Meaning	% False Detections Goal	Estimated % False Detection via CALIPSO comparison
-9	Bad or missing radar data		
5	Significant return power but likely surface clutter		
6-10	Very weak echo (detected using along-track averaging)	< 50 %	44 %
20	Weak echo (detection may be artifact of spatial correlation)	< 16%	5 %
30	Good echo	< 2 %	4.3 %
40	Strong echo	< 0.2 %	0.6 %

Table 1 – Description of CloudSat cloud mask values, false detection rates, and percentage of false detections. The percent of false detection is given by 100 times the number of false detections divided by the total number of detections for the specified cloud mask value.

Users are cautioned that radar detections with cloud mask values between 6-10 contain large numbers of false detection. These possible detections represent hydrometeors whose radar-reflecivity is below the single column sensitivity limit of the radar (about –30 dBZe), and have only been identified as a result of an aggressive along track averaging algorithm. For most applications, users should consider using cloud-mask values of 30 to 40, which are high-confidence detections.

Figure 10. Description of the CloudSat cloud mask values, false detections goals and estimates, and warnings on the use of the values (From NASA 2007a).

CloudSat has difficulty detecting some low level stratus, cumulus, non-drizzling stratocumulus, warm altocumulus composed of small water droplets, and optically thin, high cirrus (Mace et al. 2007). Information on how CloudSat attempts to overcome these deficiencies can be found in Mace et al. (2007). To account for the low level deficiency of CloudSat, we removed all data at the

lowest levels (1 km) of the cloud mask column from consideration during CloudSat mask data retrieval. WWMCA is also known to have difficulty detecting high cirrus, and we made no modifications to the CloudSat to address this issue.

D. DATA SET

1. World Wide Merged Cloud Analysis

The WWMCA data set used for this study contained hourly global cloud amount analyses for 2010 and was provided by the USAF's 16th Weather Squadron in ASCII format. Each data file provided consisted of hourly analyses for each month for approximately 1.5 million global data points. A sample of the data format can be seen in Figure 11. Data extracted from these files for this study included: year, month, day, hour, WWMCA I and J coordinates, total cloud amount from all layers, and pixel age. Further information on how we determined whether an individual WWMCA cell had enough CloudSat pixels to warrant combining for analysis is located in Section F.2.

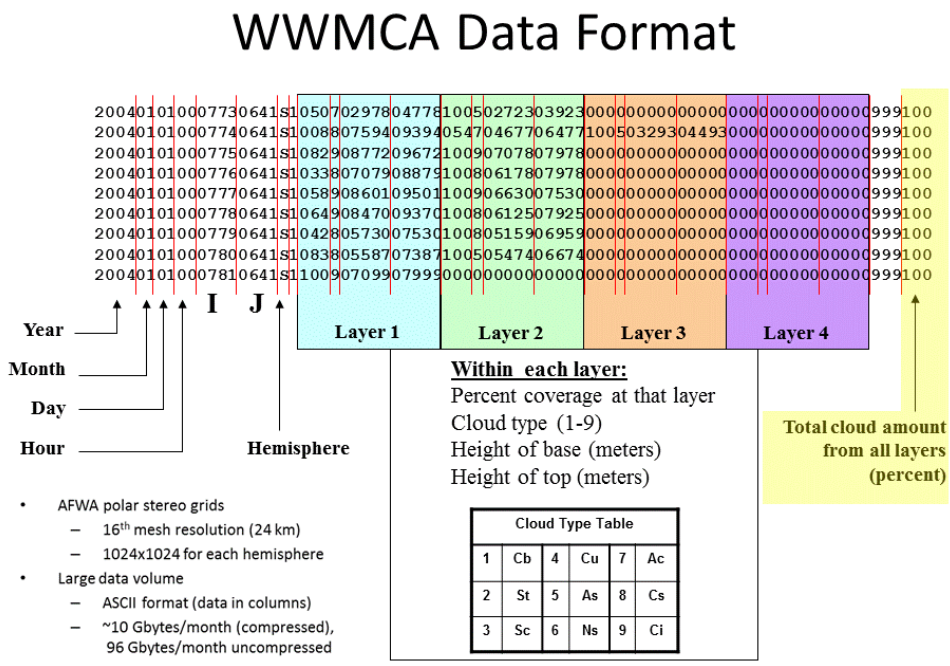


Figure 11. WWMCA data format. Cloud amount data was determined from the total cloud amount from all the layers, not from layer percentages. (From Cleary 2012b).

2. CloudSat–2B GeoProf

The CloudSat 2B–GeoProf data set used for this study was downloaded via file transfer protocol (FTP) from the CloudSat Data Processing Center in hierarchical data format (HDF). We created a data file for each orbit of the CloudSat and included detailed information for approximately 37,000 data points. CloudSat data provided mostly continuous coverage for the year 2010, with the exception of a complete outage of data from 01–15 Jan 2010 and an occasional single orbit outage throughout the rest of the year. Data extracted from these files included: date, time (hh:mm:ss), latitude, longitude, CPR Cloud Mask maximum mask value (0–40), the level of occurrence of the maximum mask value (1–124), MODIS cloud confidence flag (0–3), and MODIS cloud fraction.

E. DATA PREPARATION

1. Overview

In order to do a comparison of WWMCA analyses to CloudSat, we had to conduct a three dimensional time and space matching of the two data sets. The steps to ensure proper matching of data sets included: (1) retrieval of required data from CloudSat granules; (2) merging of CloudSat pixels from within the CloudSat granules into WWMCA compatible data sets for time and location; (3) reduction of the CloudSat data to a reasonable "truth" window comparable to the WWMCA valid time; and (4) matching the CloudSat data to corresponding WWMCA data via valid time, date, and location. Once dimensional matching was completed, a data format quality control was performed, and the data was decoded and reformatted for calculation using MATLAB. We used MATLAB for format and data quality control, reformatting, and data processing to ensure the quality of the data and assimilate the data into predetermined bins for analysis. MATLAB coding samples for all steps can be found in Appendix B.

2. Data Retrieval from CloudSat

During this step, data was retrieved at two time intervals to match with the WWMCA valid time (VT). This included 30 and 15 minute intervals prior to the top of the hour. Samples of 30 and 15 minute data sets were then processed through the final reduction steps. After a comparative review of the results, we determined that the 15 minutes would be the optimal time in order to maximize useable data and to analyze data that would be most representative of the actual conditions at the top of the hour. Thus, the CloudSat data was reduced to include only data in the Northern Hemisphere that occurred 15 min prior to the top of each hour for the raw data described in Section C.1. Details and results on the process of time determination can be seen in section G.1 of this chapter. In this step, CloudSat pixel latitude and longitude were converted into WWMCA box and WWMCA cell coordinates. Additionally, a CloudSat occurrence value was calculated to relate the CloudSat cloud mask value for each pixel to a cloud

occurrence value according to Table 3. This data was placed into CSV files for each day sorted by a time stamp. An example of the output of this phase can be seen in Figure 11.

Table 3. Translation from CloudSat cloud data mask to occurrence values.

Cloud Mask Value	**Cloud Occurrence Value**	**Interpretation of Occurrence Values**
-9	999	Missing or bad data. Will be removed from further processing
0-10	0	Clear
20	1	Probable Cloud, will be considered cloud if using 20 Threshold, will be no cloud if using 30 threshold
30-40	2	Cloud

Date	Cloudsat time	Cloudsat Lat	Cloudsat Lon	WWMCA I	WWMCA J	WWMCA box	Cloudsat Max Mask	Level of Max Cloudsat value	Cloudsat Occurrence Value ("-1" "0" "1" "2" or "999")	MODIS_cloud_flag	MODIS_cloud_fraction
20100116	21:45:17	31.4524899	-124.725044	315	711	43	40	85	2	0	0
20100116	21:45:17	31.4621029	-124.727615	315	711	43	40	86	2	0	0
20100116	21:45:17	31.471714	-124.730179	315	711	43	20	100	1	2	0
20100116	21:45:17	31.4813251	-124.73275	315	711	43	10	83	0	3	100
20100116	21:45:17	31.4909382	-124.735321	315	711	43	10	84	0	3	100
20100116	21:45:18	31.5005493	-124.737892	315	711	43	8	67	0	3	100
20100116	21:45:18	31.5101624	-124.740456	315	711	43	8	84	0	3	100
20100116	21:45:18	31.5197735	-124.743027	315	711	43	5	101	0	3	100
20100116	21:45:18	31.5293846	-124.745598	315	710	43	5	101	0	3	100
20100116	21:45:18	31.5389977	-124.748169	315	710	43	5	101	0	3	100
20100116	21:45:18	31.5486088	-124.75074	315	710	43	6	32	0	3	100
20100116	21:45:19	31.5582199	-124.753311	315	710	43	6	32	0	3	93
20100116	21:45:19	31.5678329	-124.75589	315	710	43	6	32	0	3	100
20100116	21:45:19	31.5774441	-124.758461	315	710	43	10	34	0	1	18
20100116	21:45:19	31.5870552	-124.761032	315	710	43	10	33	0	0	0
20100116	21:45:19	31.5966663	-124.763611	315	710	43	10	33	0	0	0
20100116	21:45:19	31.6062775	-124.766182	316	710	43	10	34	0	0	0
20100116	21:45:20	31.6158905	-124.768753	316	710	43	10	34	0	0	0
20100116	21:45:20	31.6255016	-124.771332	316	710	43	10	34	0	0	0
20100116	21:45:20	31.6351128	-124.773911	316	710	43	10	99	0	0	0
20100116	21:45:20	31.6447239	-124.776482	316	710	43	10	99	0	0	0
20100116	21:45:20	31.654335	-124.77906	316	710	43	10	99	0	0	0
20100116	21:45:20	31.6639462	-124.781631	316	710	43	10	99	0	0	0
20100116	21:45:21	31.6735573	-124.78421	316	710	43	30	99	2	0	0
20100116	21:45:21	31.6831684	-124.786789	316	710	43	8	75	0	0	0

Figure 12. A sample of the initial data pulled and calculated in phase I. Note that this sample has data for three different WWMCA cells.

3. Merging CloudSat Granules and Calculating Additional Information

In this step, CloudSat pixels that were located in the same WWMCA cell and contained the same date and hourly time stamp were merged into a single line of data, in order to calculate a cloud fraction for average CloudSat cloud cover from the cloud occurrence values for that cell and time. The percent cloudiness and final cloud category was calculated twice using the cloud mask threshold value of 20 as cloud and as not a cloud. An example of the output of this phase can be seen in Figure 13.

WWMCA Valid Time	Average Lat	WWMCA i	WWMCA j	WWMCA Box	WWMCA Box Length	WWMCA Box Area (km**2)	WWMCA Box Coverage (%)	Night Flag (D=day;N =night)	50 Degree (A=above; B=below)	Total CSPs	# CSPs Cloudy	# CSPs PC	# CSPs No Cloud	Sum CSPs	CS % Cloudy (cats 1 and 2)	Cloud Cover Bin (cats 1 and 2)	CS % Cloudy (cat 2)	Cover Bin (cat 2) (0-19:0 ; 20-79:1 ; 80-100:2)
2010011622z	31.424	315	711	43	19.41422199	376.9120157	7.96737667	D	B	21	15	1	5	21	76.19	1	71.43	1
2010011622z	31.563	315	710	43	19.44069122	377.940475	3.026931688	D	B	8	0	0	8	8	0	0	0	0
2010011622z	31.664	316	710	43	19.45983281	378.6850929	4.909092105	D	B	13	1	0	12	13	7.69	0	7.69	0

Figure 13. A sample of the merging of CloudSat data for the initial data shown in Figure 12. Note that instead of 42 lines, the sample now contains only three lines, one representing each WWMCA cell for a specific date and time.

4. Matching CloudSat and WWMCA via Time, Date and Location

Once CloudSat data for a specific WWMCA cell and time was reduced to a single line, it then was matched to the corresponding WWMCA data for the same location and time. After the data was matched, then only the required WWMCA information needed was retrieved from the WWMCA files. This data was placed into a single CSV file where each line had a combination of the necessary CloudSat and WWMCA data for each WWMCA cell and valid time. At this point all data was matched for all 12 months and data processing for comparisons could be started. An example of the output of this phase can be seen in Figure 14. This step was completed for all months using the 15 minute window and one day using the 30 minute window for future validation of the temporal data reduction choice made for CloudSat.

WWMCA Valid Time	Average Lat	WWMCA i	WWMCA j	WWMCA Box	WWMCA Box Length	WWMCA Box Area (km**2)	WWMCA Box Coverage (%)	Night Flag (D=day;N =night)	50 Degree (A=above ;B=below)	Total CSPs	# CSPs Cloudy	# CSPs PC	# CSPs No Cloud	Sum CSPs	CS % Cloudy (cats 1 and 2)	Cover Bin (cats 1 and 2) (0-19:0 ; 20-79:1 ; 80-100:2)	CS % Cloudy (cat 2)	CS Cloud Cover Bin (cat 2) (0-19:0 ; 20-79:1 ; 80-100:2)	WWMCA Cloud Cover Bin (0;1;2)	WWMCA Total Cloud Cover (%)	WWMCA Pixel Age (min)
201001162	31.424	315	711	43	19.41422	376.912	7.967377	D	B	21	15	1	5	21	76.19	1	71.43	1	2	100	32
201001162	31.563	315	710	43	19.44069	377.9405	3.026932	D	B	8	0	0	8	8	0	0	0	0	0	0	32
201001162	31.664	316	710	43	19.45983	378.6851	4.909092	D	B	13	1	0	12	13	7.69	0	7.69	0	0	0	32

Figure 14. A sample of a final data set after the merging between CloudSat and WWWMCA data sets. This is a continuation of the sample shown in Figures 12 and 13.

5. Data File Format Quality Control and Decoding

This step ensured that all lines from the final monthly .CSV files created during data reduction and assimilation were the same length (i.e., had inputs for all data sections), and that there were no errors in the output formatting. In this step we also converted the data to a .mat file for use with MATLAB. If any lines differed in length, they were identified and then omitted from further processing. Next, the quality controlled .mat file was decoded and the columns and variables were associated to respective categories for quick reference use in further MATLAB coding.

F. DATA CALCULATIONS AND FUNCTIONS

a. MATLAB Step 3a: Function to Calculate

In this step, we calculated WWMCA performance metrics using standard 3x3 and 2x2 contingency tables (Tables 4–5) and verification metrics. We calculated the performance by cloud category (Cloudy, Partly Cloudy, and Clear), region, and month. We used the 3x3 contingency tables for each region to assess overall hits, and one category and two category differences (Table 4). We used the 2x2 contingency tables to separately assess the performance for each of the three cloud categories (Table 5). We used as our performance verification metrics: hit proportion, miss proportion, level-1 and level-2 differences, proportion correct (PC), threat score (TS), false alarm ratio (FAR), probability of detection (POD), probability of false detection (POFD), Heike skill score (HSS), and bias (B) (Wilks 2006). These metrics are most commonly applied to the verification of forecasts. However, for our study, we applied them to the verification of analyses --- specifically, WMMCA analyses. In the following descriptions of these metrics, the letters A-D are referred to and represent the quantities described in Table 5.

Table 4. A generic 3x3 contingency table for assessing WWMCA hits, and level-1 and level-2 differences.

		Observations (CloudSat)		
		Cloud	Partly Cloudy	Clear
Analysis (WWMCA)	Cloud	HIT	MISS – Level-1 difference	MISS – Level-2 difference
	Partly Cloudy	MISS – Level-1 difference	HIT	MISS – Level-1 difference
	Clear	MISS – Level-2 difference	MISS – Level-1 difference	HIT

Table 5. A generic 2X2 contingency table for assessing WWMCA performance in analyzing the occurrence of the three cloud categories: Cloudy, Partly Cloudy, and Clear. The example shown in this figure is for the analysis of Cloudy conditions. We used similar tables for assessing WWMCA performance in analyzing Clear and Partly Cloudy conditions (After Wilks (2006) and Cleary (2012)).

		Observations (CloudSat)			
		Cloud	Not-Cloudy		
Analysis (WWMCA)	Cloudy	A Hits (A / N)	B False Alarms (B / N)	A + B (A + B) / N	Marginal Totals for Analysis & Marginal Distributions for Analysis
	Not-Cloudy	C Misses (C / N)	D Correct Rejections (D / N)	C + D (C + D) / N	
		A + C (A + C) / N	B + D (B + D) / N	N = A + B + C + D 1.00	
		Marginal Totals for Observations & Marginal Distributions for Observations		Sample Size & Total	

A = Number of Hits
B = Number of False Alarms
C = Number of Misses
D = Number of Correct Rejections

N = A + B + C + D
Letter/ N = Probability of that event

In the 3x3 matrix, when cloud categories agreed this was a hit. When they did not agree this was a miss. When a miss occurred and the cloud category differed by one category this was considered a level-1 difference. When a miss occurred and the cloud category differed by two categories, this was considered a level-2 difference. The total number of hits were added together and divided by the total sample size to get an overall measure of success referred to as the hit proportion (HP). The total number of misses was added together and divided by the total sample size to determine the miss proportion (MP).

Before evaluating performance metrics it is useful to know the distribution of the observations to know how often an event had the opportunity to be evaluated. It is also useful to know the distribution of the analyses for comparison to the distribution of the observations. Mathematically, the marginal distribution (MD) for observations and analyses are calculated by:

$$\text{MD OBSERVATIONS} = (A+C) / N$$

$$\text{MD ANALYSES} = (A+B) / N$$

The proportion correct (PC) is the ratio of correct analysis of an event to the total number of samples and is an accuracy measurement. Accuracy measures reflect the correspondence between pairs of forecasts and the events they are meant to predict (Wilks 2006). Even though the PC does not clearly distinguish between correctly identified event and non-event occurrences, Wilks (2006) considers this one of the most straightforward and sensitive measures of the accuracy of non-probabilistic forecasts for discrete events. The PC credits yes and no events equally, and thus can be problematic when the yes event is rare. In our study, we found that WWMCA rarely detected partly cloudy conditions, so the PC may not be the best assessment of skill for this particular

cloud category, but it may work well for verifying the skill of cloudy or clear condition forecasts. Mathematically, the PC is calculated by:

$$PC = (A + D)/N$$

The threat score (TS), also known as the critical success index (CSI), provides information on the accuracy of the performance by eliminating the times when a correct rejection was identified. This score is useful when the event that is forecasted occurs substantially less frequently than non-events occur. A value of one indicates the best possible threat score while a value of zero indicates the worst (Wilks 2006). TS is calculated by:

$$TS = CSI = A / (A + B + C)$$

The false alarm ratio (FAR) is a reliability performance metric. It is the fraction of yes forecasts that turn out to be incorrect or the proportion of forecasts that never materialize. Due to its negative orientation, a smaller FAR is preferred with the best performance a value of zero and the worst being one. Often this performance metric is called the false alarm rate; however the term false alarm rate is actually reserved for the discrimination measurement of probability of false detection (POFD) (Wilks 2006). Mathematically, FAR is calculated by:

$$FAR = B / (A + B)$$

The POFD is also known as the false alarm rate (F). It is one of two discrimination performance metrics, the other being the POD. It is the ratio of false alarms to the total number of nonoccurrences of the event. It provides

the conditional relative frequency of a wrong forecast given that the event does not occur (Wilks 2006). Mathematically, F is calculated by:

$$POFD = F = B / (B+D)$$

The POD, often also referred to as the hit rate, is the ratio of correct analyses of the event to the number of times that the event occurred. This performance metric is only concerned with the actual event of interest and its occurrence. Combined with the POFD, this metric provides the conceptual and geometric basis for the signal detection approach for verifying probabilistic forecasts (Wilks 2006). Mathematically, POD is calculated by:

$$POD = \text{Hit Rate} = A / (A + C)$$

The HSS provides an overall measure of the skill of the WWMCA analyses. Perfect WWMCA analyses would result in a HSS of one, while values between zero and one would indicate improvement over random analyses, and negative values would indicate skill worse than random analyses. The HSS provides a measure of how skillful the analyses were compared to random analyses (Wilks 2006). This is probably the best single metric to use for performance assessment because it: (1) describes the analysis skill by comparisons to a random analysis benchmark; and (2) gives more (less) credit to accurately analyzing rare (common) events. Mathematically, HSS is calculated by:

$$HSS = 2((A * D) – (B * C)) / ((A + C) (C + D) + (A + B) (B + D))$$

The bias is an indicator of the over or under forecasting of an event. It is calculated by combining the values for hits and false alarms and dividing them by the number of observations of the event. Unbiased analyses have a bias of 1. Bias values greater than 1 indicate an over analysis of the event meaning the event was analyzed more often than observed, while values less than 1 indicate an under analysis of the even or that the event was forecast less often than observed (Wilks 2006). Mathematically, bias is calculated by:

$$\text{Bias} = (A + BA) / (A + C)$$

b. MATLAB Step 3b: Main Program to Calculate

This step was critical for mass data analysis. This took the core function formulas and allowed them to be rapidly calculated for various geographic regions and CloudSat thresholds. This program can be used to create a more detailed reduction thresholds as determined by the user. This step is where the reduction for WWMCA cell coverage and the various summated time categories and regions of interest were input and then performance metrics calculated. The calculations were put into a matrix for easy reference during this step, although there is no associated labeling of the rows and columns. As the focus of the research was revised, codes for step 3a and 3b were modified to reflect these revisions.

2. Final Conversion for Analysis: Rewriting to CSV

In this step, the .mat files were converted to monthly .csv files for integration back into Microsoft Excel. Once in .csv format, files were opened and the proper titling of the rows and columns was inserted for identification.

G. DATA ANALYSIS

Most calculations were completed though the use of MATLAB but needed to be displayed via a different method to properly analyze and compare the

statistics. Microsoft Excel proved to be the simplest way to do side by side comparisons between months, geographic regions, and cloud categories. To best do this, the twelve monthly files that were the output of MATLAB were condensed into a single Excel file consisting of ~6,400 rows of data for all months and valid times. In addition to the contingency table metrics described in Section F, we also determined the percent of monthly and annual data from each geographic region in relation to the total data set. This information can be found in Appendix C. Once all raw data was compiled into a single file, a combination of line and bar graphs was used to analyze the data.

Before a detailed study could be completed for all locations, months, and time periods, we needed to determine which CloudSat data period to use to represent the truth, how much data coverage overlap between WWMCA and CloudSat would be sufficient for a proper analysis, and which CloudSat data mask would provide the most accurate representation of the clouds that occurred.

1. CloudSat Period Determination

We needed to determine a proper time period to average CloudSat data for our "truth" to be representative of the top of the hour analyses. It is important to note that clouds can change dramatically within an hour, and thus we needed a short enough time window to not be greatly impacted by the evolution of cloud elements within a WWMCA cell. The choice of the time period for the "truth" value for this study was motivated by concerns that a longer period could lead to more inaccurate results about what occurred at the top of the hour, and a shorter period could lead to too little data and not cover all regions we wanted to investigate. Additionally, we were concerned with cloud advection and wanted to minimize the risk of significant cloud advection through a WWMCA cell during the period of truth. We investigated using CloudSat data from 15, 30, and 60 minutes prior to the hour for which a top of the hour WWMCA analysis is valid. It was necessary to balance a time period that allows for enough data coverage to

establish a valid study, but not so long a time period as to dilute the averaged "truth" that is to represent the top of the hour cloud conditions over an area. Prior studies used 30 and 60 min averaging of CloudSat.

A hypothetical example (Figure 15) illustrates the impacts of using different averaging periods. Using a 60 minute averaging of cloud would lead to the top of the hour representation of truth to be LT 20%, or Clear, when at the top of the hour it was actually Cloudy. Using a 30 min window, the average would be 20-80%, or Partly Cloudy. Using the 15 min window nearest the top of the hour, the average would be greater than 80%, or Cloud. As the averaging time decreases, the average cloud conditions become more representative of the conditions at the WWMCA valid time. Based on these considerations, we decided that the 15 minute window provided a sufficiently large data set for our study, while ensuring a fair representation of the top of the hour cloud coverage.

Time period	Cloud Cover
12:00-12:30	CLEAR
12:30-12:40	FEW
12:40-12:50	SCT
12:50-13:00	OVERCAST

Time avg	Conditions Reported
60 min avg	LT 20%
30 min avg	20-80%
15 min avg	GTE 80%

Figure 15. Observed conditions over a location for one hour and the conditions that would be reported by CloudSat as the "truth" for the top of the hour based on various averaging periods.

2. Coverage Area Determination

In addition to determining a "time window" that would be representative of the top of the hour; we needed to determine how much CloudSat coverage was needed to describe the actual cloud conditions in a WWMCA cell. Due to the

nature of the size and structure of clouds and cloud layers, it is unreasonable to expect one single 1.3 km x 1.7 km CloudSat pixel to be representative of an entire 12 km x12 km or 24 km x 24 km WWMCA cell. If this were done in an area with varying clouds there would be large number of possible cloud representations, each of which could be different than the observed conditions. Thus, we decided that we needed to determine the minimum acceptable areal coverage of a WWMCA cell by CloudSat pixels to allow an assessment of WWMCA performance in that cell. In prior studies (e.g., UCAR 2008, AER 2010, Cleary 2012) data coverage determination for CloudSat to WWMCA comparisons was done via a CloudSat percentage method. In this method, the average WWMCA cell size was used and the number of pixels a CloudSat granule could have in that cell was calculated based on the WWMCA cell dimensions at 60 degrees latitude. Next, the number of pixels that did occur in the WWMCA cell was summed up and if that cell included "6 or more CloudSat points" which was approximately 25% of the possible CloudSat points, then sufficient CloudSat data was determined to be available to assess WWMCA performance in that cell. See Figure 16 for a pictorial representation of this process.

If there were sufficient data amounts, then a cloud percent was calculated for CloudSat as follows:

$$\frac{Number\ of\ Cloudsat\ cloudy\ points}{Total\ number\ of\ Cloudsat\ points} * 100 = Total\ cloud\ amount\ (\%)$$

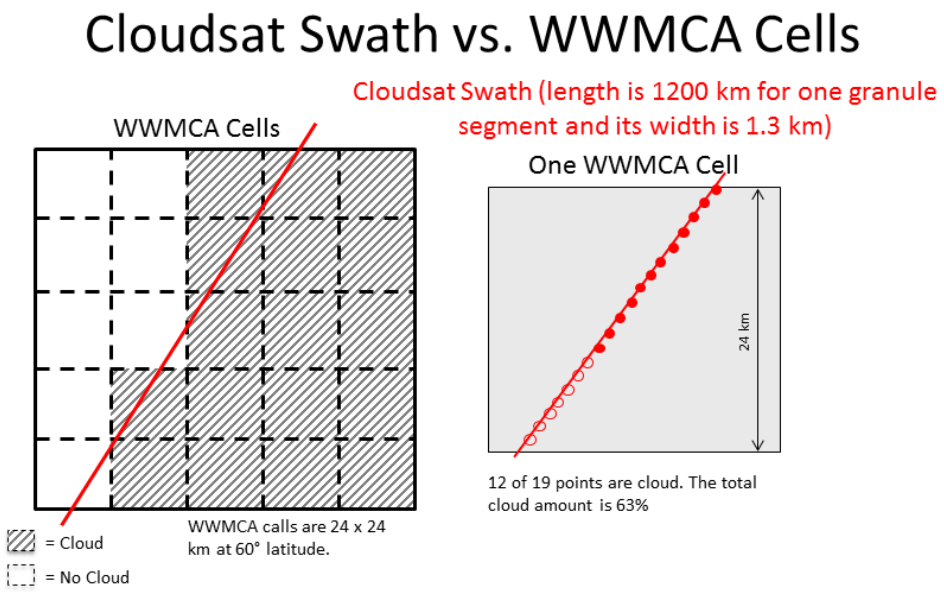


Figure 16. CloudSat Swath vs WWMCA Cell at 60 degrees latitude (From Cleary 2012b).

Further research into this method demonstrated that since WWMCA cells vary by latitude, even though six CloudSat pixels may represent 25% of the possible CloudSat pixels in a cell at 60 degree latitude, it only represents 1.8% of a WWMCA cell at 90 degrees latitude and 7.2 % of a WWMCA cell at zero degrees latitude (See Table 6). In our study, we worked to improve the determination of whether there was sufficient CloudSat data for a given cell by: (a) analyzing cell size by latitude; (b) calculating the percent area of each WWMCA cell; and (c) calculating the percentage of the maximum possible coverage of the cell represented by CloudSat pixels. From this information, we determined an acceptable cell coverage amount and assessed WWMCA performance only for cells and times for which the percent coverage threshold was met.

Table 6. A comparison by latitude of WWMCA cell resolution, maximum CloudSat pixels per cell, the percent cell coverage the maximum pixels represents, and the minimum cell coverage available using a 6 pixel minimum requirement. Note the ranges between 9 and 19% for maximum cell coverage and from 1.8% to 7.2% using 6 pixels.

Latitude	Resolution	Max CloudSat Pixels	% cell coverage	% cell coverage with 6 CloudSat pixels
0N	12.5 km	16	~19.0%	~7.2%
60N	24 km	30	~9.0%	~2.0%
90N	25 km	32	~9.5%	~1.8%

Figure 17 shows CloudSat coverage by latitude in a full dataset for March 2010. Note that there is a great variation by latitude in maximum possible coverage, ranging between 7% and 14%. Some months saw a potential maximum of 19% cell coverage. We needed to represent all latitudes, so we considered 6% data coverage a good threshold for minimum coverage.

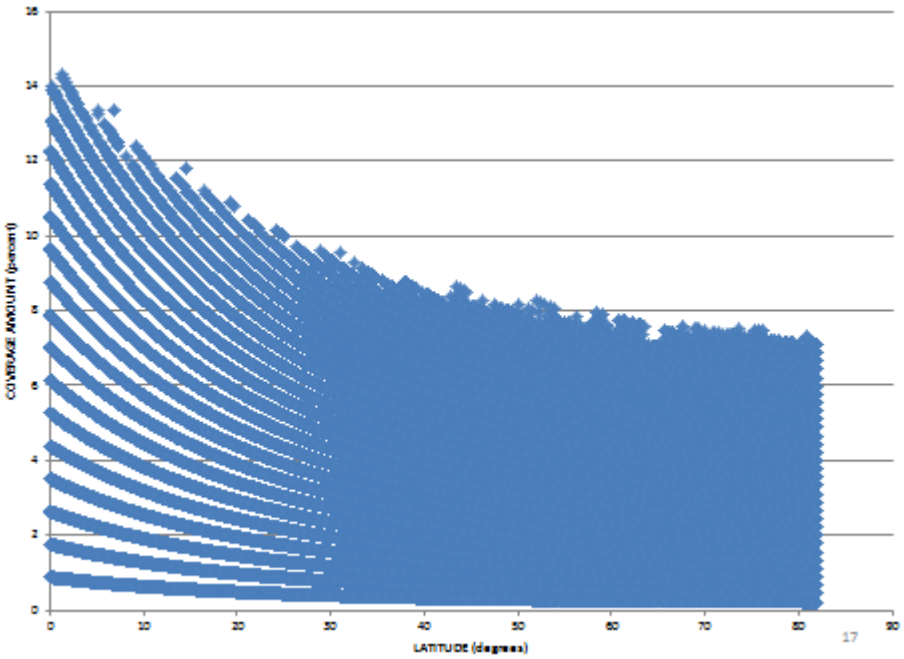


Figure 17. Percent coverage of WWMCA cells by CloudSat, by latitude for March 2010.

We were interested in how results might vary if we used a smaller coverage amount, so we conducted a comparison over the Northern Hemisphere for January 2010 to assess the impact of limiting the data source to only data from a coverage area greater than 6%. We compared data sets with no limitations in area coverage to data sets that had the greater than 6% WWMCA cell coverage limitation. We noticed that using all cells that had at least one matching CloudSat pixel resulted in a much greater decrease in performance (Figure 18) versus the 6% or greater coverage. This could be due to the fact that a 1.3 x 1.7 km square is not a fair assessment of cloud conditions for a 12 km x 12 km or 24 km x 24 km cell. While 6% area coverage may not be a lot either, it is significantly larger and can be considered a better evaluation. We determined 6% would be our threshold value of minimum WWMCA cell coverage to maximize CloudSat and WWMCA overlap area while ensuring the entire hemisphere could be analyzed. Due to the coverage area variation by latitude,

we feel that future approaches to WWMCA verification should consider using a latitudinal based approach to data coverage versus a CloudSat pixel percentage.

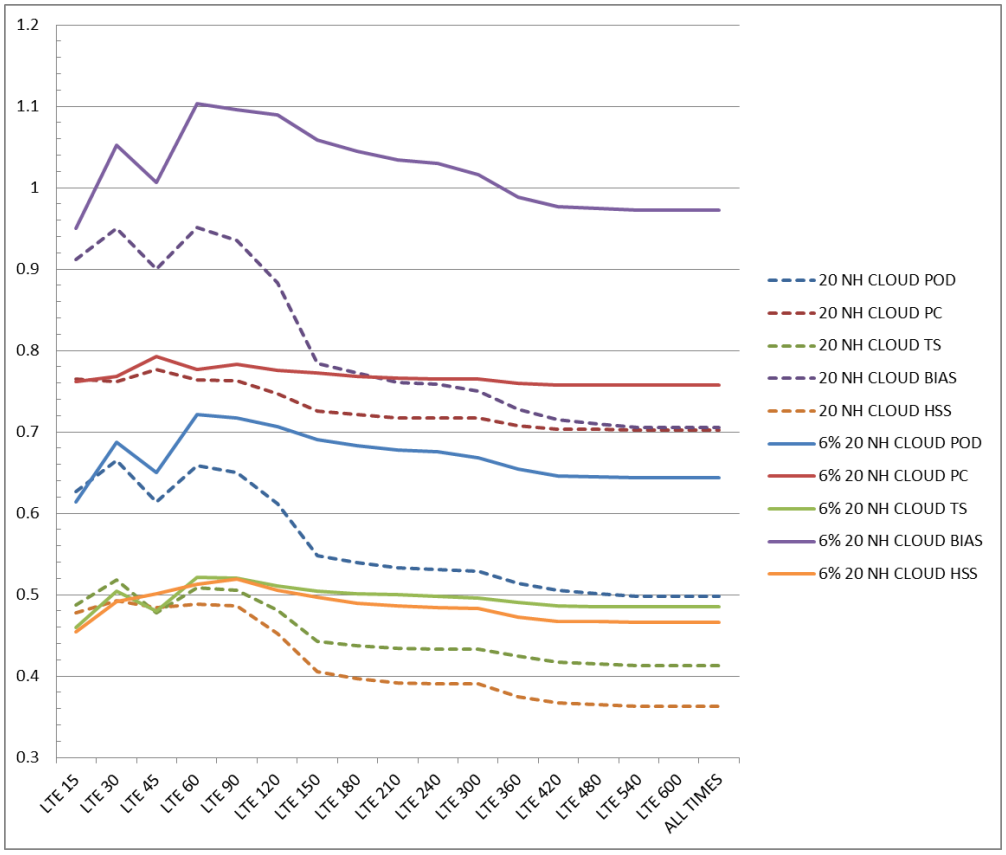


Figure 18. Comparisons of NH performance metrics for January 2010 from using 100% of CloudSat data available (dashed lines) and using CloudSat data only when it covered at least 6 % of a WWMCA cell (solid lines). Notice the significant decrease in skill when using all available CloudSat data.

3. CloudSat Data Mask Threshold Determination

A comparison between the CloudSat cloud mask values of 20 and 30 for cloud thresholds was conducted before our final geographic analysis could be completed. UCAR (2008) and AER (2010) used 20 as the lowest value to indicate cloud for their research, while Cleary (2012) used both 20 and 30 as thresholds and determined that WWMCA performance was rather insensitive to the choice of these thresholds in representing cloud. We decided that we should

also investigate this threshold. A comparison of performance metrics for the cloud category using both 20 and 30 values as the minimum indicator for cloud is seen in Figure 19. On average, there is a less than 2% difference in the various performance metrics for the Northern Hemisphere when using the separate values as a cloud threshold. We compared our findings with information from the NASA CloudSat documents, and with results from prior studies, and decided to use the threshold of 20 to indicate when a cloud occurred for a CloudSat pixel. Note that if our use of this 20 threshold leads to errors in our analyses of WWMCA performance, these errors will tend to underscore WWMCA performance.

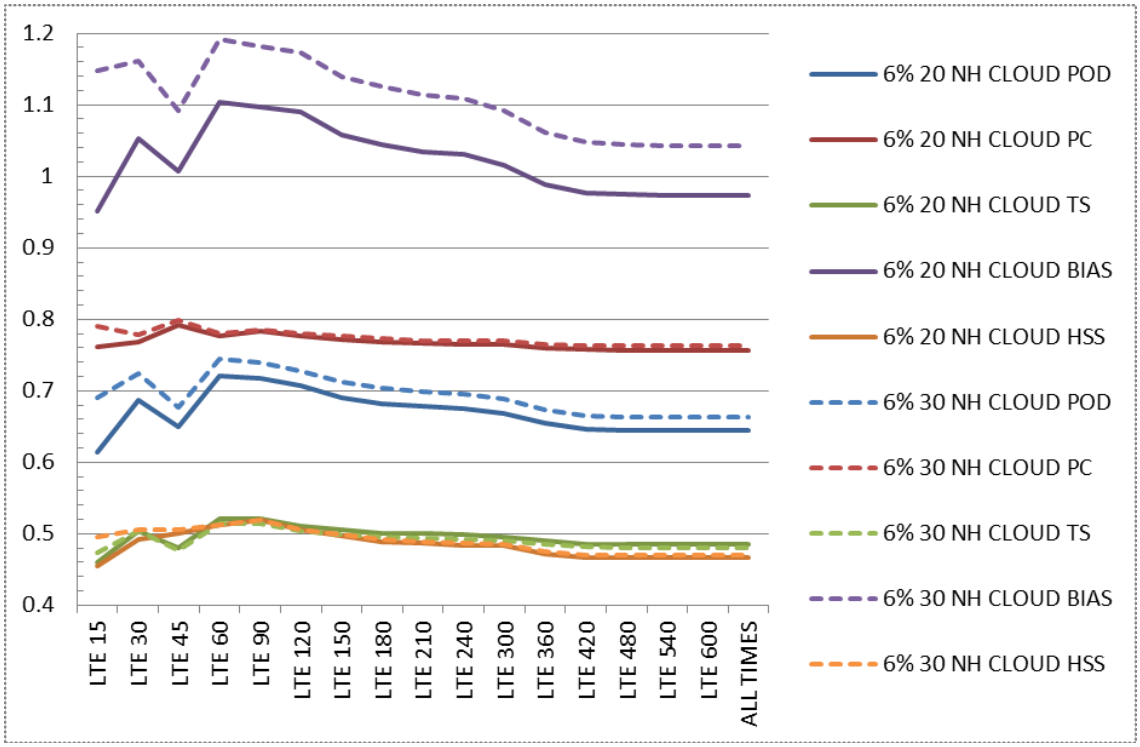


Figure 19. Comparisons of the CloudSat Cloud mask thresholds of 20 and 30 for the Northern Hemisphere for five performance metrics. The solid lines represent the 20 threshold while the dashed lines represent the 30 threshold.

H. ASSUMPTIONS AND APPROXIMATIONS

After reducing the data set based on the results from our preliminary investigations, the remaining performance metrics were calculated, graphed, and analyzed for various latitudinal bands and locations, cloud categories, and input data ages. There were eight geographic regions of interest for this study into which we binned our data: six latitudinal bands (0N–90N, 0N–50N, 0N–23.5N, 23.5–35N, 35N–50N, and 50N–90N) and two smaller geographic regions (Southwest Asia [WWMCA Box 12]; and South China Sea [WWMCA Box 22]). Our assessments of WMMCA performance were the main focus of this study and these results are discussed in Chapter III.

We used the following assumptions and approximations to make our study feasible:

1. CloudSat data describes the actual cloud conditions.
2. The highest cloud mask level reported on the CloudSat profiler is a good measure of overall could cover.
3. Cloud cover amounts are accurately represented by three cloud categories: Cloudy, Partly Cloudy, and Clear.
4. A 15 minute window prior to the top of the hour is a good estimate of cloud cover for the top of the hour.
5. Two CloudSat passes per day over a given area are representative of the entire day for that area.
6. CloudSat data that covers at least 6% of the area of a WWMCA box is an acceptable minimum coverage requirement.
7. Clouds are indicated by CloudSat if the mask value is 20 or greater.
8. WWMCA performance can be adequately assessed using metrics based on 2x2 and 3x3 contingency tables.

9. The four main seasons can be represented by three calendar months: winter (January, February, and March), spring (April, May, and June), summer (July, August, and September) and autumn (October, November, and December).

I. IMPROVEMENTS OVER PRIOR METHODS

The combination of the study length, hemispheric size, inclusion of time latency issues, and the range of performance metrics evaluated for this study provide increased benefits over prior studies. As noted in Chapter II, most studies considered only a relatively short study period for a relatively large study region, or a small study region. Very few studies considered both long study periods and large study areas. Those that did, focused on only a couple of performance metrics. By conducting an annual study on the entire Northern Hemisphere with various subcategories and performance metrics, we were able to provide more detailed information on all aspects of WWMCA. The following list of items denotes specific improvements over prior studies:

1. Year-Long Study

Our study considered all 12 months while, with the exception of the Stubblefield, all other studies covered between 10 days and four months. This increased not only the size of the data sets, but allowed for comprehensive comparison between months, seasons, and annual performance versus treating one season or month as representative of all months and seasons.

2. Increased Number of Performance metrics Evaluated

Our study addressed 11 different performance metrics for evaluating WWMCA performance. UCAR (2008), AER (2010), Norquist (2007), and Ruggiero (2000) only concerned themselves with hit rate while Horseman (2007) added miss rate. Heidemen (1995) expanded from hit rate to look at level-2 differences and Gustafson (2011) considered differences greater than 20% between WWMCA and his truth. Cleary (2012) used more extensive measures

of performance metrics than others that are similar to our study but focused on smaller geographic regions for location verification.

3. Hemispheric and Latitude Banded Approach

Our study investigated WWMCA data on various levels: hemispheric, latitudinal bands, and regional. Only Gustafson (20110, UCAR (2008), and AER (2010) have conducted large scale studies beyond a specific region. These were done on a global and hemispheric average. Norquist (2007) and Horseman (2007) limited themselves to one or a few WWMCA cells, while others kept regions limited to two to six WWMCA boxes. Our study will allow for comparison to other hemisphere studies and to Cleary's (2012) regional study. Our study will also help determine if there is variation in performance between latitude, an issue that has only been briefly addressed in prior studies.

4. Monthly and Seasonal Comparisons

We have provided an evaluation of monthly, seasonal, and annual performance that has only been attempted by Stubblefield (2011) and Cleary (2012). The inclusion of all 12 months and seasons has the potential to provide improved insights to help modify tuning algorithms compared to what can be determined from shorter period studies.

5. Use of Three Cloud Categories

Our study addressed overall performance, and performance in three specific cloud categories: Clear (<20% cloud cover), Partly Cloudy (20-80% cloud cover) and Cloudy (>80% cloud cover). Studies such as Horseman (2007), UCAR (2008), and AER (2010) only considered the overall hit/miss of WWMCA, not the performance for individual cloud categories.

Furthermore, different values have been used to determine if conditions were cloudy or clear. UCAR (2008) and AER (2010) used the threshold of 1% cloud as differentiation between clear and cloudy skies, Stubblefield (2011) used 30% as the threshold between cloudy and clear, and the threshold for Horseman

(2007) was not provided. All studies measured hit rate; however, it is hard to make direct comparisons of the results due to the varying thresholds used for cloud conditions.

6. Improved Area Overlap in WWMCA and CloudSat for Baseline Data Assessment

UCAR (2008), AER (2010), and Cleary (2012) used a six CloudSat pixel limit to determine if there was enough data for comparative analysis for a WWMCA cell. Our data processing included a latitudinal adjustment and considered the amount of the WWMCA cell covered, not just a percentage of the amount of possible CloudSat pixels that could have occurred. This serves to improve the comparison of results between different latitudes as well as enhance the quality of data for the study as the percentage of a grid cell actually sampled is more standardized.

7. Assessment of the Impacts of Input Data Age

Unlike any prior studies, we assessed the impacts on WMMCA performance of varying the age of the data that was used by WWMCA. This included assessing the impacts on performance: (a) as older data was incorporated into the data set; and (b) by starting with old data and then including newer data. This assessment of the impacts of data time latency on WMMCA performance also helps establish time performance curves for use in modifying input algorithms and threshold cutoffs.

J. LIMITATIONS OF THIS STUDY

Our methods provided some improvements over prior methods; however, our study was limited by the multidimensionality of the topics being addressed, and by the quantity, quality, and complexities in processing the available data.

CloudSat is in a 1331L and 0131L orbit, and thus is limited in its ability to represent variations throughout a day.

Using 6% and greater coverage of the WWMCA cell, is an improvement over the 6 pixel standard used in some prior studies; however, this still amounts to very little coverage in comparisons to the size of the WWMCA cells. Using CloudSat alone as truth, the maximum WWMCA cell coverage possible is only 19%, and this is just at the tropics.

We assessed WMMCA performance in six latitudinal bands and two sub-regions, which allowed us to be more specific in our assessments than would be possible using only global or hemispheric regions. There could still be important performance variations within our regions that our assessments would not be able to identify (e.g., variations due to differences in the background surfaces, terrain, climatological conditions).

We did retain the flag for separating day and night data; however, we did not evaluate and address the separation of daytime and nighttime performance of WWMCA.

We attempted to remove concerns in accuracy due to backscatter and miss readings at the lowest 1 km from CloudSat by removing data from the levels that occur in this range. However, we only evaluated the WWMCA overall cloud percentage and did not evaluate the four reported layers of WWMCA. Because of this, there may be periods in which WWMCA reported clouds at this level that could be considered misses in this study, if they were in a shallow layer that did not extend above 1 km and there were no other clouds above them.

We assessed WMMCA performance a monthly and annual scales, and established baseline performance metrics unavailable from prior studies. However, we did not attempt to determine how well WWMCA performs for specific short term weather phenomena (e.g., blocking conditions with persistently clears skies, large scale weather systems, monsoon patterns, localized weather patterns).

.

III. RESULTS

A. OVERVIEW

The focus of this study was to provide a base line evaluation of the performance, or operational health, of WWMCA. The results shown in this chapter are based on the 2010 WWMCA data set for the Northern Hemisphere and the corresponding CloudSat data set as described in Chapter II, and using the processing and analysis methods described in Chapter II (e.g., using CloudSat data from 15 minutes prior to the hour, a CloudSat data mask threshold of 20, and a CloudSat data minimum coverage threshold of 6% of the corresponding WWMCA cell). Section B of this chapter describes the general performance of WWMCA based on hit and miss proportions, as well as level-1 and level 2 differences, for six latitudinal bands (0N-90N, 0N-50N, 0N-23.5N, 23.5-35N, 35N-50N, and 50N-90N). Section C describes how WWMCA performance varied with the timeliness (or latency) of the input data by latitude bands. Finally, Section D describes the performance of WWMCA for four metrics in analyzing the three cloud categories for two latitude bands and for two specific geographic regions, Southwest Asia and the South China Sea. Additional results are provided in Appendix C.

B. GENERAL PERFORMANCE

The hit proportion, miss proportion, and level-1 and level 2 differences were used to evaluate the overall ability of WWMCA to accurately identify cloud conditions at annual and monthly scales and for the six latitudinal bands. Figure 20 shows the Northern Hemispheric annual hit proportion was 63%. The miss proportion was 37%, and level-1 differences occurred in 24% of the cases while level-2 differences occurred in 12% of the cases. Performance was best near the equator and worsened as latitude increased, and was especially low in the highest latitude band (50N-90N). The performance in the highest latitude band

was worse than the hemispheric average and was 20-28% lower than for the other bands.

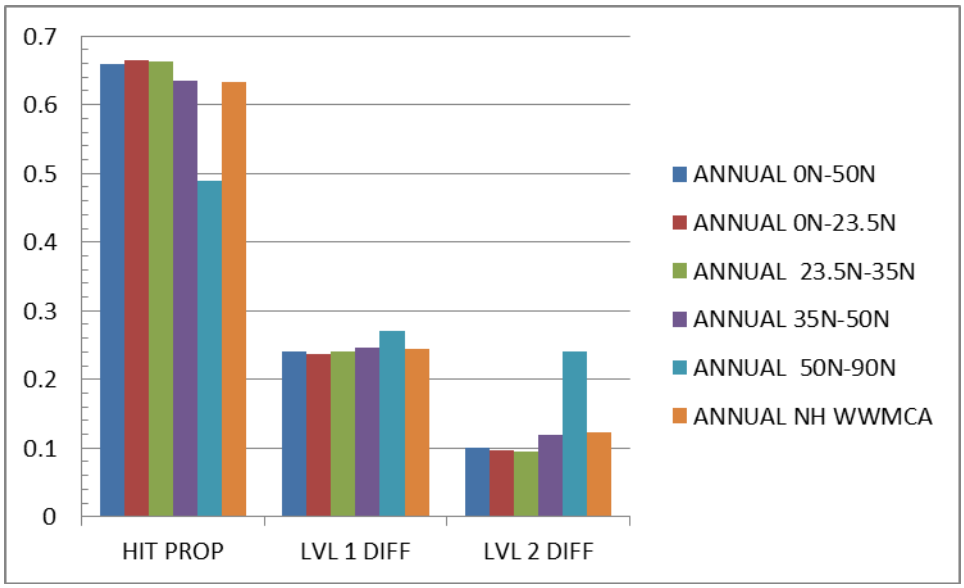


Figure 20. Annual mean results for hit proportion and level-1 and level-2 differences for the six latitudinal bands.

Table 7 summarizes the monthly performance ranges by band (see also Figures 21-22). Note that the best (worst) performance occurred at low (high) latitudes, and that there was little month to month variation in performance within each band. The monthly mean versions of the annual mean results (Figure 20) are shown in Appendix C.

Table 7. Monthly ranges for hit proportion, and in level-1 and level-2 differences, for each latitude band. Note that the best (worst) performance occurred at low (high) latitudes, and that there was little month to month variation in performance within each band.

Region	Hit proportion monthly range	Level-1 difference monthly range	Level-2 difference monthly range
0N – 50N	64% – 67%	22% – 25%	8% – 12 %
0N – 23.5N	62% – 68%	22% – 25%	7% – 12%
23.5N – 35N	61% – 69%	21% – 26%	6% – 12%
35N – 50N	60% – 66%	22% – 28%	9% – 16%
50N – 90 N	42% – 56%	23% – 29%	16% – 32%
0N – 90N	62% – 65%	23% – 26%	11% – 14%

Figure 21 and Figure 22 display the monthly comparisons for each region. The top panel in Figure 21 shows overall WWMCA performance in the Northern Hemisphere. Note that the detection of cloud categories was slightly better near the solstices and that the level-2 hit error was the worst in the fall and winter. The 0N-50N latitude band was chosen for specific investigation, as it represents the coverage provided by geostationary satellites that refresh every 15–30 minutes, although this band does have a few areas that depend on polar orbiters. The middle panel in Figure 21 displays data for this region. The best overall performance for this band occurred in the winter followed by the solstices. On average there was a less than 10% level-2 difference for this area with the minimum level-2 difference seen in the late fall and winter months.

The 50N-90N latitude band is covered by only polar orbing satellites due to parallax in the geostationary satellites. Polar satellites provide coverage for approximately 18 hours of the day, and provide updates to CDFS II on average every 75–90 minutes, due to constraints imposed by orbit and data downlink

capabilities. As a result, the highest time latencies observed in WWMCA usually occur in the 50N-90N band. The bottom panel in Figure 21 displays data for this region and shows that the hit proportion was up to 20% lower than the hemispheric average depending on the month. This overall decrease in performance may well be linked to the latency of the data in this band. In addition, there is a well-defined late spring to summer period of maximum hit proportions and a well-defined minimum in the fall. The 50N-90N band has the largest deviations between the maximum and minimum performance, which could be due to the seasonal changes in the background. The level-2 difference for the high latitudes is 5–15% larger than for any other region, with a minimum difference in the summer months. This minimum deviation suggests that cloud detection may be better during periods when there is less ice and snow in the backgrounds. It is known that snow and ice are difficult backgrounds to distinguish from cloud and the trend in error rates, specifically the level-2 difference, for the high latitude regions tends to support this statement.

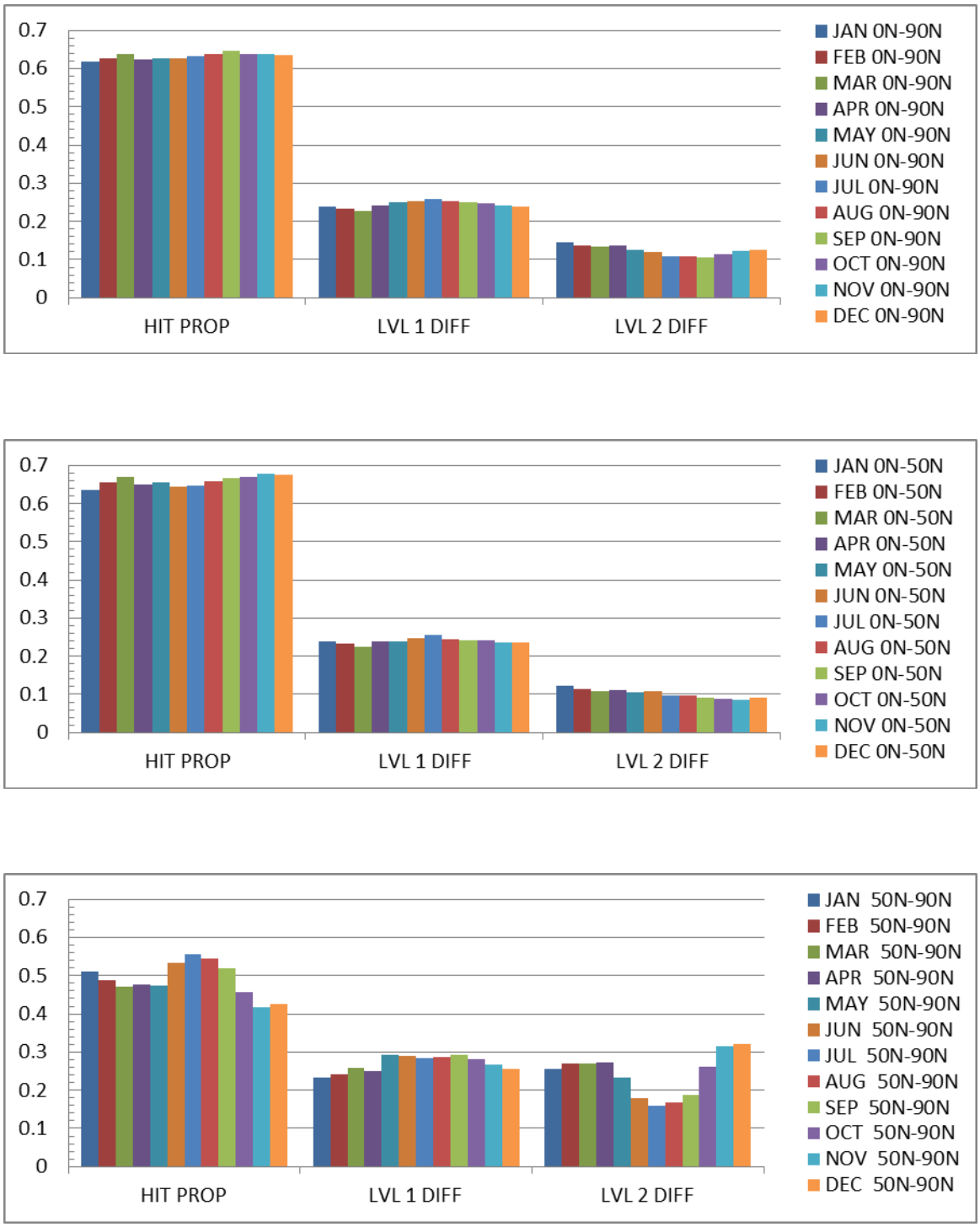


Figure 21. Monthly results for hit proportion, and level-1 and level-2 differences, by latitudinal bands. From top to bottom, the chart represents the following regions (0N–90N, 0N–50N, and 50N–90N). Note the different patterns of performance over the various regions.

The best performance for accurate cloud category detection was seen in the lower latitudes. The top panel in Figure 22 displays data for the 0N–23.N region while the middle panel shows 23.5N–35N. No defined seasonal trends appeared which may be due to the fairly consistent background conditions throughout the year. These tropical and subtropical bands had the lowest level-2 differences of all the latitude bands. The high performance in these regions may be due to a combination of factors, such as more frequent satellite refresh rate, consistent backgrounds with little variation between seasons, large uniform backgrounds (e.g., ocean surface), or the relatively slowly varying major cloud patterns (e.g., the inter-tropical convergence zone (ITCZ); relatively persistent clear skies in much of the subtropics).

The middle latitude band covered 35N–50N. Many studies consider mid-latitudes to extend up to 60N; however, we limited our band to 50N to account for the current maximum range of geostationary satellite coverage. Like the 0–50N band, this band is a blend between geostationary and polar data due to 'cusp' regions between the geostationary satellite coverage areas (Figure 1). The results for this region can be seen in the bottom panel of Figure 22. No specific trends in the seasons were noted for this region; however the overall performance was lower than for the lower latitudes. This region is a transition region between the geostationary region to the south and the polar obiter dominated region to the north. This may explain the decrease in accuracy, as dependency on polar orbiters increases where geostationary coverage is lacking.

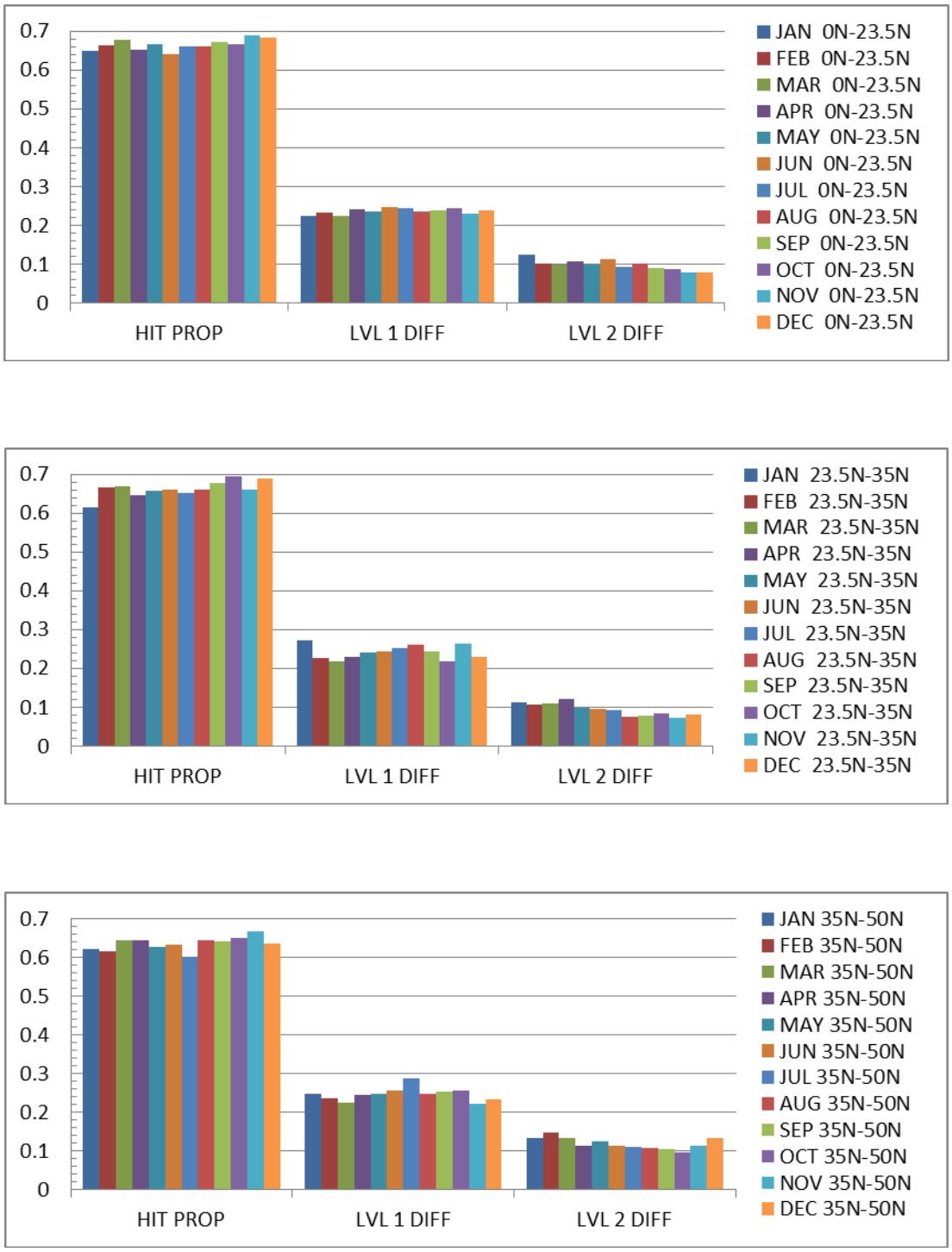


Figure 22. Monthly results for percent hit, and level-1 and level-2 differences, by latitudinal bands. From top to bottom, the chart represents the following regions (0N–23.5N, 23.5N–35N, and 35N–50N). Note the different patterns of performance over the various regions.

The overall results shown in Figures 21-22 indicate that WWMCA performed best in low latitudes and worst in the high latitudes. At lower latitudes, there was generally less than 5% variation between months. But variation up to 15% can be seen in higher latitude bands.

We also compared the results in Figures 21-22 to those from prior studies. Hit rates for April and May can be compared to the studies done by UCAR (2008) and AER (2010), while September and June results can be compared to the study by Gustafson (2011). All of these studies used data from 2010. Hit proportions for our study were around 62% for April and May, while UCAR (2008) and AER (2010) determined that WWMCA was in agreement with CloudSat 75% of the time. Our lower values may be due to how cloudy and clear were defined. UCAR (2008) and AER (2010) used 1% cloud as the threshold between cloud and no cloud. Each time we had partly cloudy conditions, and sometimes when we had clear conditions, would have been a cloud hit under the used in some prior studies, which would have led to a higher performance result. Another difference may be due to differences in the truth time period used. We used 15 minutes while both the UCAR (2008) and AER (2010) studies used one hour. Finally, results may differ due to the use of different minimum coverage thresholds. These prior studies used a threshold based on a percentage of possible CloudSat pixel versus WWMCA cell area, and did not account for latitudinal variations in WMMCA cell size.

Gustafson (2011) conducted his study on a global scale, while ours was done on the Northern hemisphere; however our findings with respect to accurately matching cloud categories were similar. We used the same three cloud categories as Gustafson. Gustafson determined that WWMCA and MODIS had a clear match for 25% of the cases, a partly cloudy match for 6%, and a cloudy match for 34% of his cases. These values lead to an overall success rate for matching cloud categories in 65% of the cases, while our Northern Hemisphere results were between 63% and 65%. We did not calculate the hit proportion for each cloud category, but this could easily be done in a follow-on

study. The slight differences in values could be because Gustafson's study was for the globe scale and ours was for the Northern Hemisphere. It could also be due to the time period used for truth averaging, as Gustafson used 60 minutes and we used 15 minutes, or that MODIS was used versus CloudSat for verification of WWMCA analyses. While MODIS is similar to CloudSat, it may be beneficial to use a combination of MODIS and CloudSat to verify WWMCA. Additionally, a comparison of MODIS to CloudSat may be beneficial to determine if one satellite better represents truth over the other.

C. LATENCY STUDY

1. Overview

Latency in WWMCA performance was investigated from two perspectives, one in which we started with older data and examined how adding younger data impacted performance, and the other in which we started with younger data and examined how adding aged data impacted performance. WWMCA starts with the previous analysis and then attempts to update each cell with the newest data available. The old analysis will persist if more recent data is not available. Neither of our two approaches replicates how CDFS II merges the data to produce WWMCA, but the combination of both should adequately determine the impacts of time latency. After our first review of the data, we determined that some time bins had too few data points in them, which led to what we determined to be spurious performance results. Because of this, the initial time bins were enlarged and time bins with less than 300 data points were removed from our study. This change had the greatest impact on the January data set, as data was only available for half of the month of January. It also removed Monthly results for percent hit, and level-1 and level-2 differences, by latitudinal bands. From top to bottom, the chart represents the following regions (0N–23.5N, 23.5N–35N, and 35N–50N). Note the different patterns of performance over the various regions very old data over the tropics, as it is rare to have data over 2 hours old for this region due to the abundance of geostationary satellite updates.

The spurious results were less noticeable after applying these restrictions, but some may still be present. Ideally, we would have been able to work with time bins containing at least 1000 data points, to ensure enough sampling took place for determining the impacts on performance of data time latency.

AFWA uses the following conditions as an assessment of the operational health of the final WWMCA product based on the timeliness of the input data. If the global average age of all WWMCA pixels is less than 90 minutes, the product is GREEN, and considered a reliable product. If this age is 90–120 minutes, the product is AMBER and should be used with caution. Finally, if the global average age is greater than 120 minutes, the product is RED and should be used with extreme caution. These conditions are based on the assumption that pixel age averaging done on a global scale is a direct reflection of the quality in the resulting cloud analysis. Addressing the impacts of latency was critical to determine if these AFWA thresholds are appropriate. With the exception of the AER (2010) study, which compared the difference in performance of WWMCA using all data versus only data less than 3 hours old, no prior studies have addressed the data latency issue.

Annual comparisons between the regions can be seen in Figure 23. For comparisons for each month between the regions to the annual hemispheric average for hit proportion, and level-1 and level-2 differences, see Appendix C. The purpose of these comparisons is to determine how WMMCA performance changes over time as data is added or taken away, versus evaluating the performance between the regions which was done in the previous section.

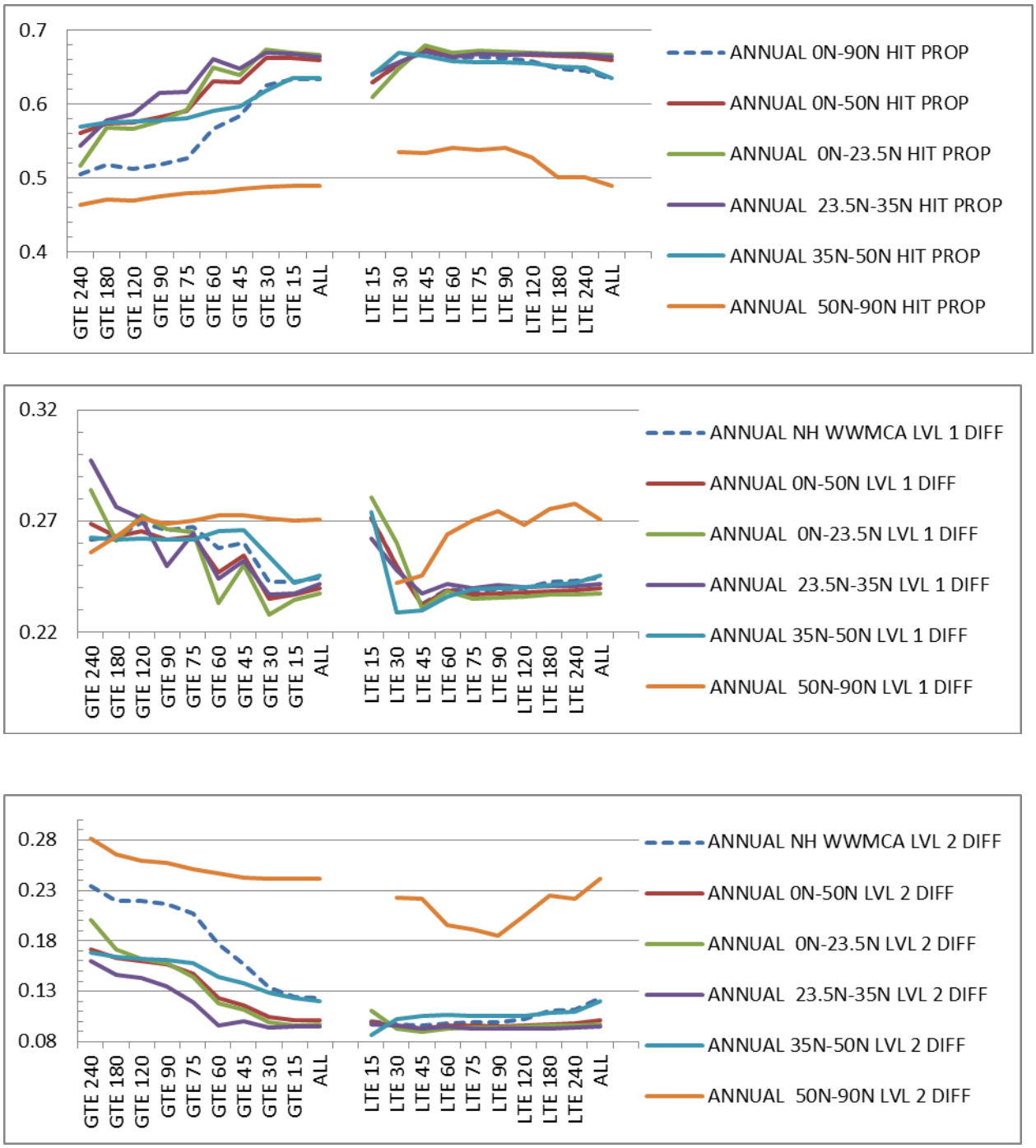


Figure 23. Annual average performance by input data age (minutes) for the six latitudinal bands. From top to bottom the panels are proportion correct, level-1 difference, and level-2 difference. Note that the y-axis has been adjusted to enhance the identification of differences between the regions. The hemispheric average is highlighted with the dashed line.

It would be beneficial to address the current process that determines how "reliable" WWMCA is. The current condition assessment looks at the global average of pixels to assess health. The rate of decline with age varies over geographic regions so a latitude based health assessment could be more beneficial to the end user. If a WWMCA target was at high latitudes it would be helpful to know that while the "global average" data age may have allowed for a GREEN interpretation of the WWMCA analyses this region's average data age is actually RED in reliability. This would allow for a more accurate assessment of the analysis and forecast due to increased knowledge of regional confidence.

2. Adding Newer Data to Old

The impacts of adding newer data to old can be seen in the left graphs in Figure 23. These graphs represent the results when starting with the oldest data and systematically adding younger data. An increase in performance as large as 20% was found from adding younger data, which supports the general assumption that using younger data improves performance over using only older data. Adding younger data to old is less relevant in the high-latitude regions than in other locations. This may be because the percentage of younger data available for the high latitudes is far less than the amount of older data, so the averaged improvements are minimized. When adding younger data to aged data, the most significant jumps occur between the 60–75 minutes period and between the 30–45 min. This is most likely an effect of the percentages of older and younger data available at these times.

3. Adding Older Data to New

CDFS II attempts to use the timeliest data available. If timely (recent) data is not available then old data from the prior analysis is persisted. This often leads to the use of very old data, especially at higher latitudes. Additionally, more timely data at a lower spatial resolution will trump slightly older data at a higher spatial resolution, which may not be an improvement in properly identifying cloud conditions. The right graphs in Figure 23 represent the impacts

on performance of adding older data to younger data. With the exception of an initial increase in performance to data ages of approximately 45 minutes, a decrease in performance is seen in all regions as older data is added. As latitude increased, the rate of decrease in performance with respect to data age increased as well. Aged data has the greatest impacts on the ability to detect clouds in the high latitudes. This does not mean that aged data performs well at low latitudes, but since aged data is rarely used at lower latitudes the impacts of small amounts of older data are minimized. Although performance decreases significantly with age, not providing any WMMCA analyses when the data is old would also be problematic and might impact customer confidence in WWMCA and CDFSII generally. The best performance occurred when the data was less than 45 minutes old. Adding additional and older data led to a decrease in performance. The hemispheric average performance did not decrease drastically, but there were changes in the amount of decrease in the latitude bands. The high latitude band had overall lower performance for all periods, but its performance was rather consistent until 90 minutes when it rapidly decreased.

D. CLOUD CATEGORY STUDY

We assessed WMMCA performance by cloud category for eight regions using seven performance metrics. Only four regions, four performance metrics, and an analysis of marginal distribution will be examined in detail in this thesis. The results for all regions and metrics are available in Appendix C. The four focus regions discussed in this chapter are 0N–50N, 50N–90N, SWA, and SCS. These regions allow for distinctions between performances for locations: (a) in which the input data is primarily from geostationary satellites or primarily polar orbiting satellites; and (b) with specific types of weather and climate. In this chapter, the focus is on assessing performance via POD, POFD, bias, and HSS.

.

1. Marginal Distribution

Marginal distribution comparisons are helpful in determining the opportunities to observe and analyze specific events. Figures 24-27 show the marginal distribution of CloudSat and WWMCA for each month and cloud category compiled by region. For the 2010 study period and the Northern Hemisphere as a whole, the CloudSat data indicated Clear for 45% of the observations, Cloudy for 38% of the observations, and Partly Cloudy for 17% of the observations. The corresponding WWMCA data indicated Clear for 49% of the analyses, Cloudy for 36% of the analyses, and Partly Cloudy for 15% of the analyses. These results show a good agreement between WWMCA and CloudSat on an annual and hemispheric average basis.

The CloudSat and WMMCA MDs are in good agreement for specific latitude bands and region, with the notable exception of the high latitude band. The MDs at the high latitudes shows WWMCA analyzed Clear most often and CloudSat observed Cloudy most often throughout the year, with differences between CloudSat and WWMCA of almost nearly 30% in the winter and fall. The smaller differences at lower latitudes indicate that WMMCA performance should be better at lower latitudes. Note that both CloudSat and WMMCA identified Partly Cloudy conditions least often. This suggests that WMMCA analyses of Partly Cloudy conditions may be problematic, due to these conditions being less common.

Gustafson determined that 65% of the time the sky was either clear or cloudy for the globe in September and June. We determined that for the northern hemisphere annual average, the sky was clear or cloudy in 83% of the CloudSat observations and in 85% of the WWMCA analyses. The corresponding numbers from our study for September and June 2010 were 82% and 85%, respectively.

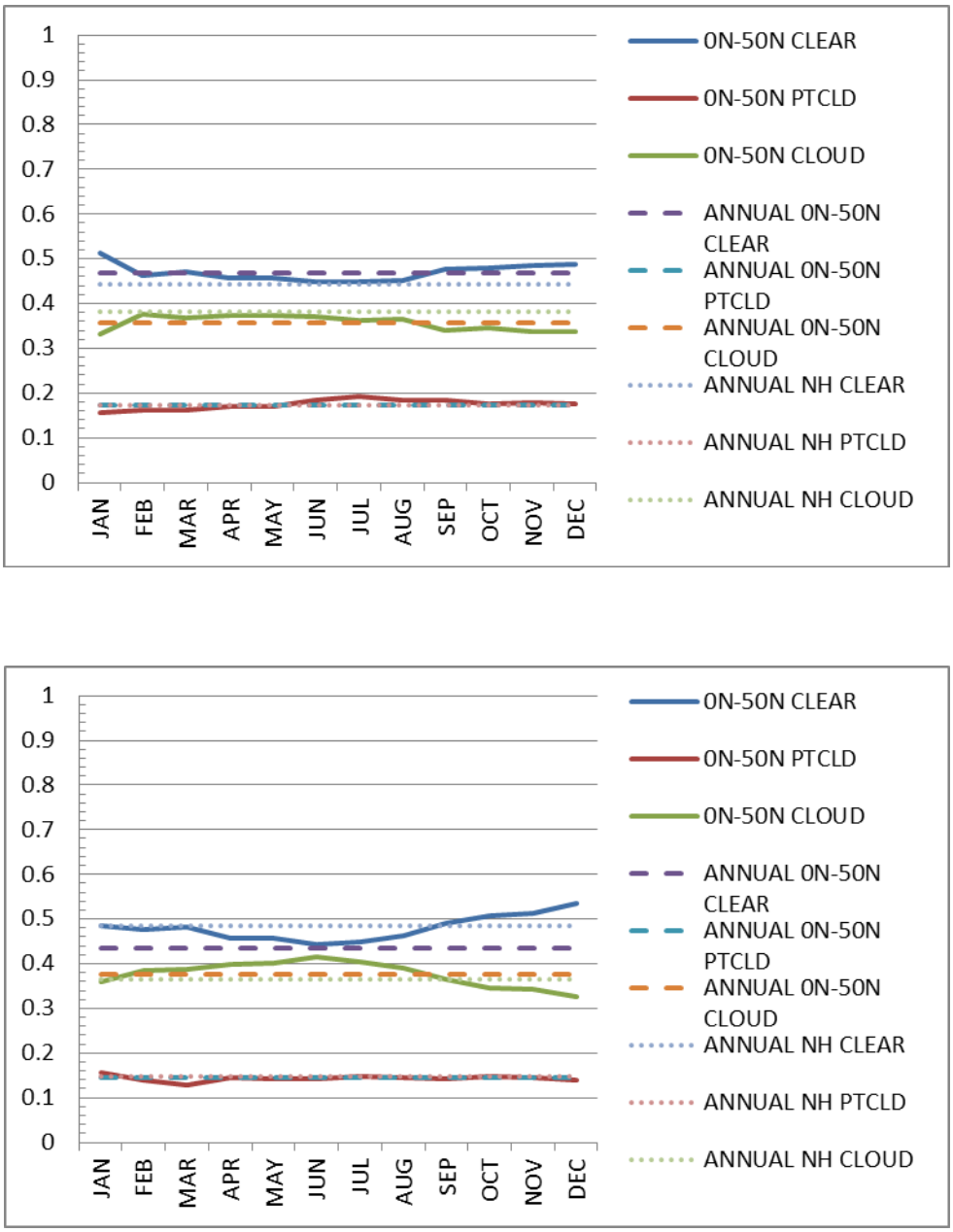


Figure 24. Marginal distributions for each month and all three cloud categories for 0N–50N. The top panel is for CloudSat while the bottom panel is for WWMCA. This figure should be compared to Figures 25–27. Solid lines represent the monthly values, while the dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

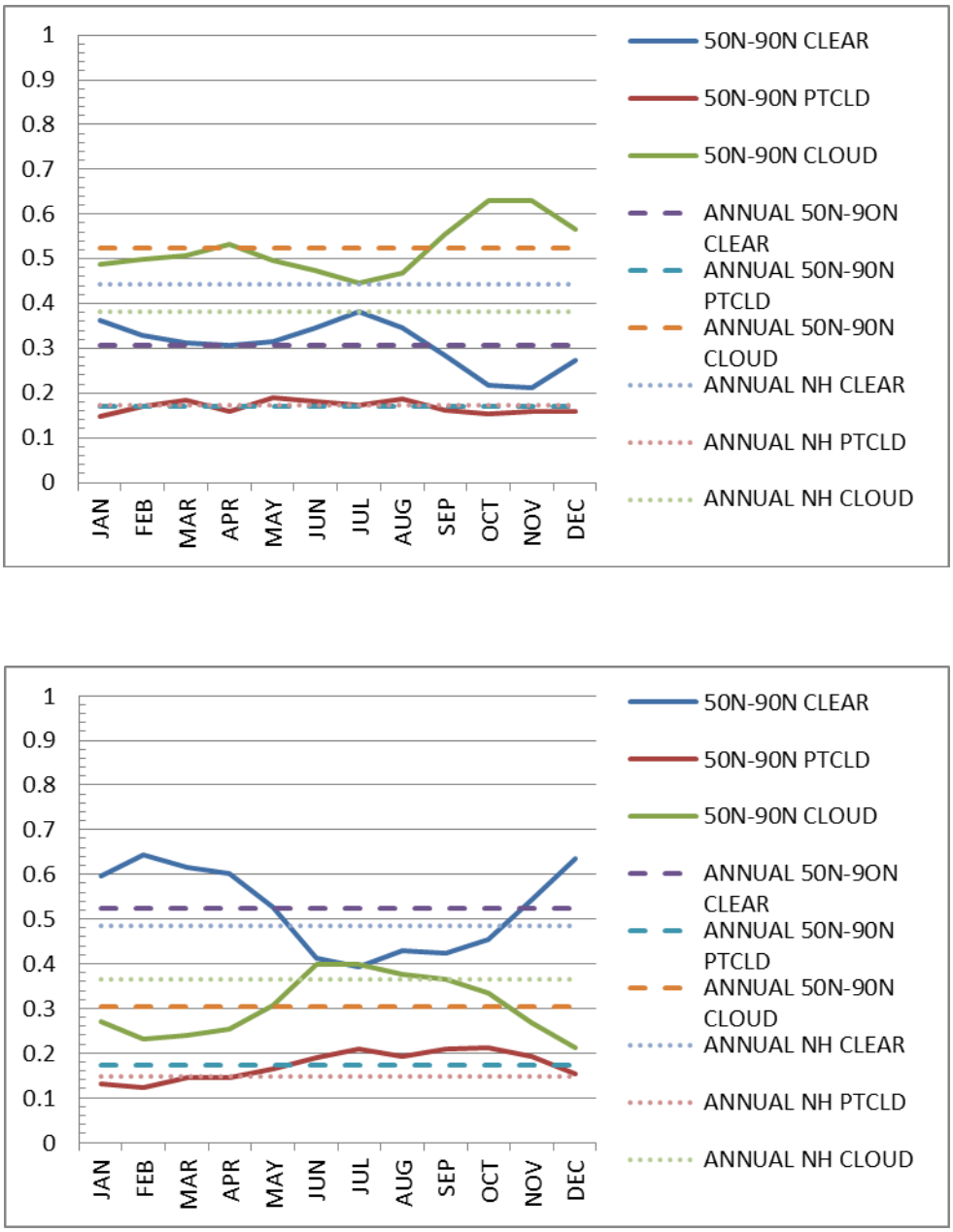


Figure 25. Marginal distributions for each month and all three cloud categories for 50N–90N. The top panel is for CloudSat while the bottom panel is for WWMCA. This figure should be compared to Figures 24, 26, and 27. Solid lines represent the monthly values, while the dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

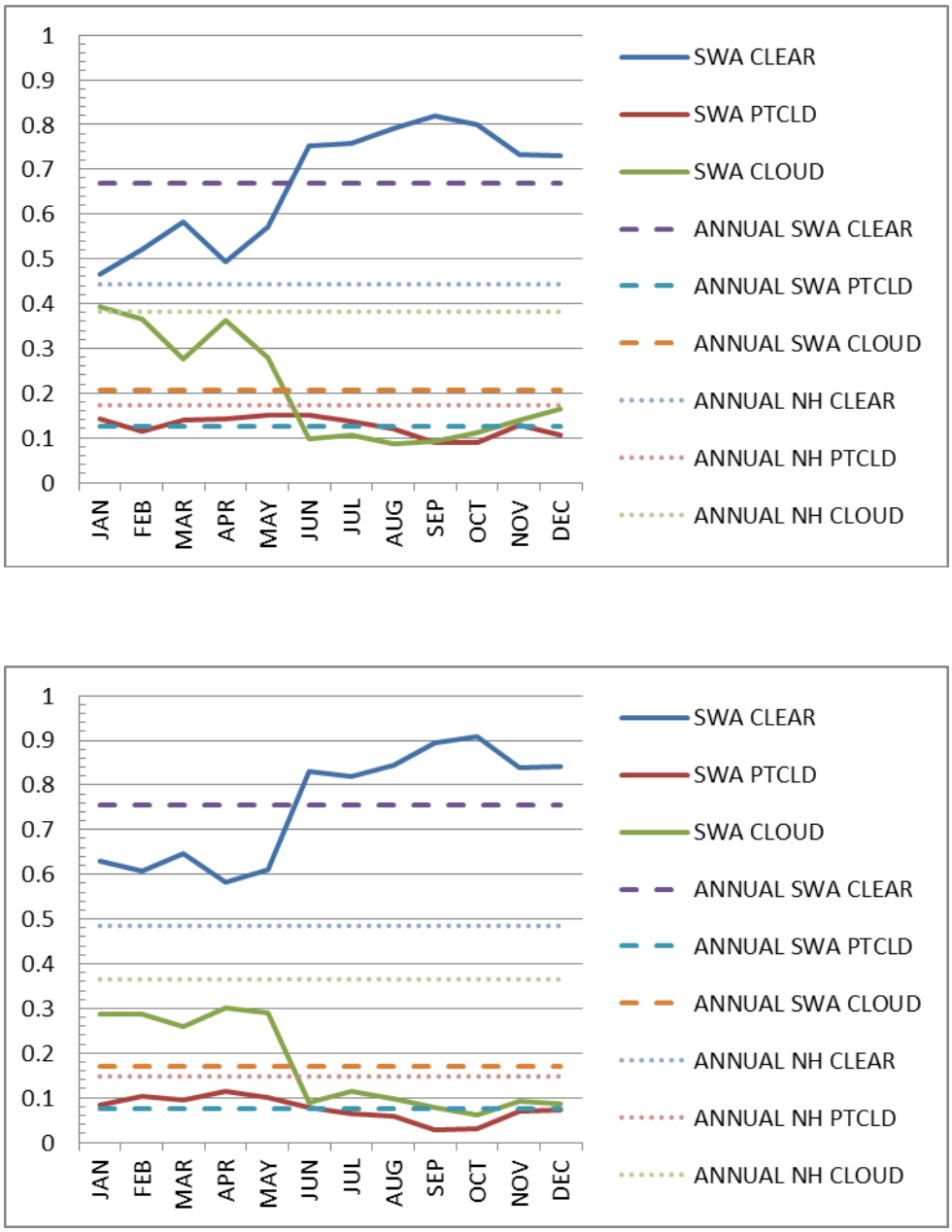


Figure 26. Marginal distributions for each month and all three cloud categories for SWA. The top panel is for CloudSat while the bottom panel is for WWMCA. This figure should be compared to Figures 24, 25, and 27. Solid lines represent the monthly values, while the dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

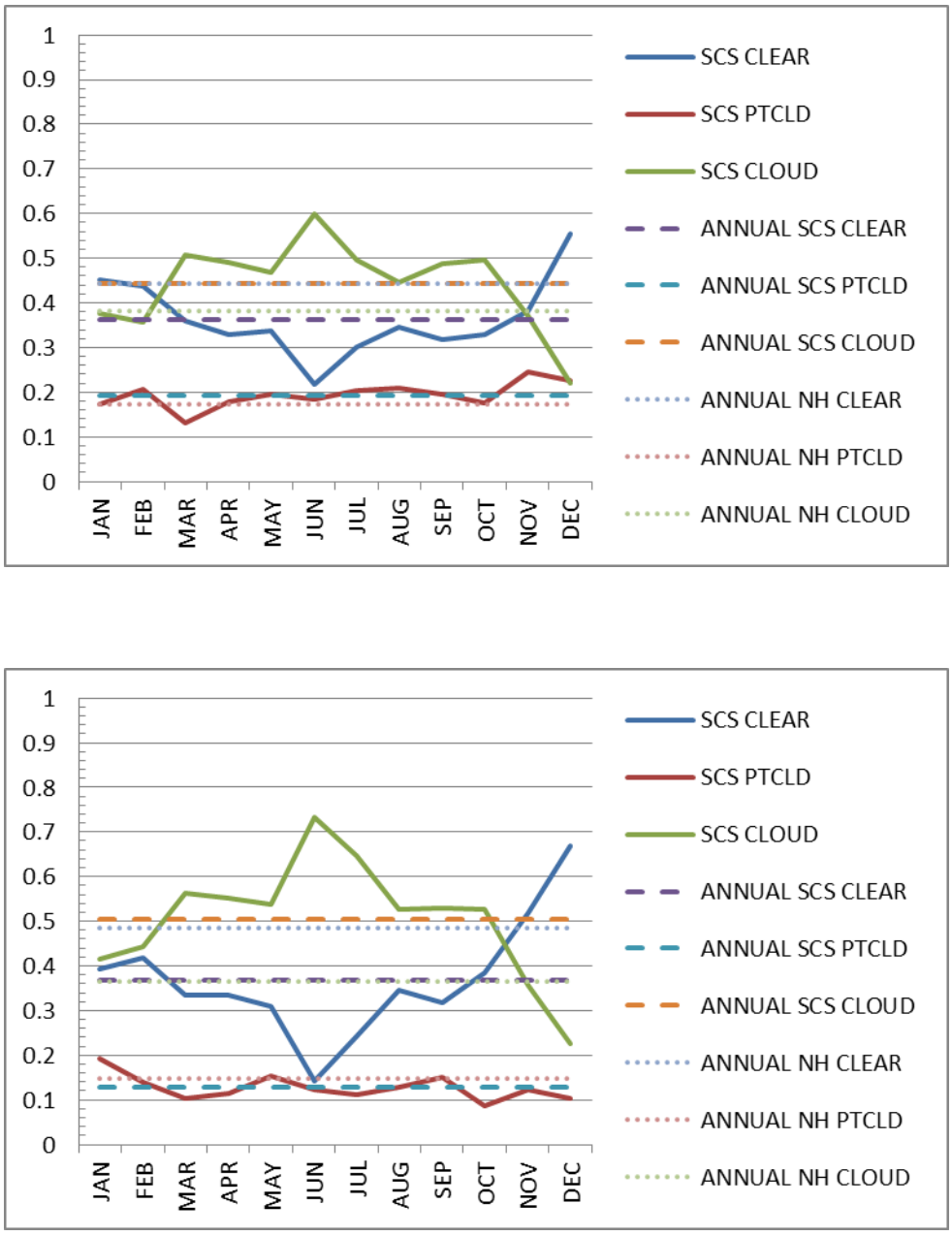


Figure 27. Marginal distributions for each month and all three cloud categories for SCS. The top panel is for CloudSat while the bottom panel is for WWMCA. This figure should be compared to Figures 24– 26. Solid lines represent the monthly values, while the dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

2. Probability of Detection

Figure 28 and Figure 29 show the probability of detection for each month and cloud category by region. The annual Northern Hemisphere POD rates were 75% for Clear, 68% for Cloudy and 22% for Partly Cloudy conditions with variations of +/- 5% for clear, +/- 2% for partly cloudy, and +/- 7% for cloudy conditions. The POD and other performance metrics for Party Cloudy conditions were routinely worse than for the two other cloud categories. This may have been due to the tendency for WWMCA to analyze predominately clear or cloudy conditions. For 0N-50N, the POD for Clear was similar to that for the NH, while the POD for Cloudy was slightly better than for the NH. For the high latitude regions, POD for Clear was best and was within +/- 10% of the NH average. However, for Cloud, POD was much worse than for the NH at the higher latitudes, and there was a greater than 30% difference between the highest (60% in July) and lowest (31% in Dec) PODs. The highest POD was for Clear in SWA, with POD as high as 95% and an annual average POD that was 15% better than the NH average (Figure 30). However, the POD for Cloudy in SWA was nearly 10% lower than the NH average. This suggests that the relatively common and persistent clear sky conditions in SWA were relatively easy for WMMCA to analyze, while the less common and transient cloudy sky conditions were more difficult for WMMCA to analyze. For the SCS, there was a reversal in the top performing cloud category, with Cloudy having the best POD and a better POD than for the NH as a whole (Figure 30). The SCS had a nearly 30% variation in POD values between seasons for Cloud and over 40% variation for Clear. This performance variation may be due to difficulty in analyzing cloud changes associated with monsoonal shifts.

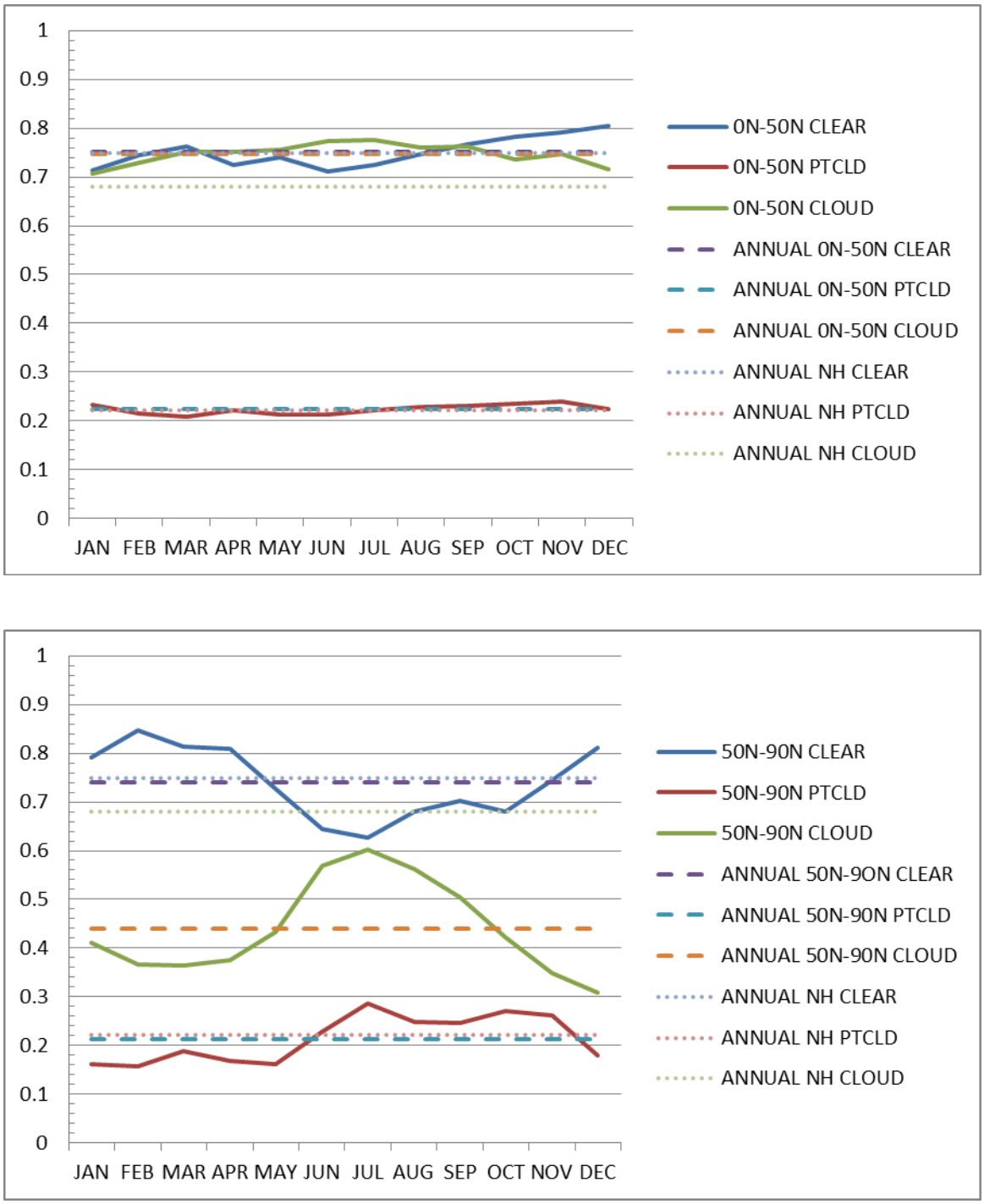


Figure 28. Probability of Detection for each month and all three cloud categories. The top panel is for 0N–50N while the bottom panel is for 50N–90N. This figure should be compared to Figure 29. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

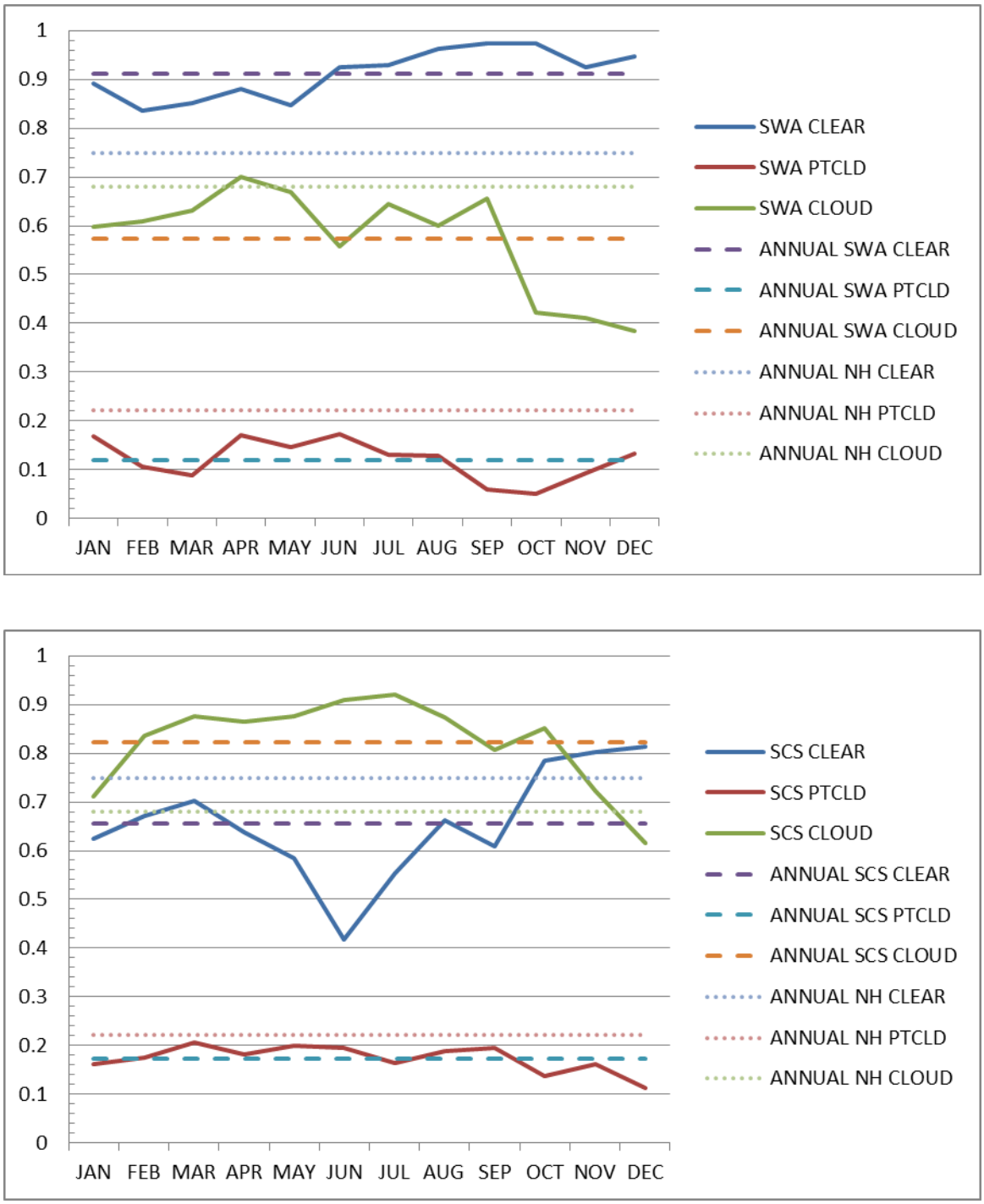


Figure 29. Probability of Detection for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figure 28. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

3. Probability of False Detection

Figure 30 and Figure 31 show the probability of false detection for each month and cloud category by region. The annual Northern Hemisphere average POFD was 27% for Clear, 17% for Cloudy, and 13% for Partly Cloudy with deviations of +/- 5% for Clear, +/- 2% for Partly Cloudy, and +/- 4% for Cloudy. The regional POFDs were similar to the NH POFDs, with the highest —that is, the worst — POFDs occurring for Clear for all regions, except for the SCS, where POFD was highest / worst for Cloudy. In general, WMMCA performance, including performance as measured by POFD, was worse for the high latitudes and better for the low latitudes. In particular, the POFD for Clear was high (48-58%) in the colder months (Nov-Apr), indicating WMMCA too often analyzed clear skies when the actual conditions were Cloudy or Partly Cloudy. However, the high latitude POFDs for Partly Cloudy and Cloudy were similar to those for the NH average. The POFD for Clear in SWA was high (40-65%) in all but four months (Feb-May), indicating that WMMCA analyzed clear skies too often. This result, when combined with those in Figure 30, indicates that WMMCA achieved a high POD for Clear by issuing too many false alarms for Clear. The POFDs for SCS were quite high for some months, especially Jun for Cloudy (99%), Aug for Partly Cloudy (69%), and Dec for Clear (49%). The monthly variations in POFD for the SCS were also quite large, with variations of more than 80% from late spring - early summer to late fall. This is most likely due to WMMCA difficulties in representing monsoon variations in clouds over the SCS, especially variations associated with the onset of the summer monsoon and intraseasonal variations due to Madden-Julian Oscillations, fluctuations in tropical cyclone activity, etc. Note that the poor performance indicated by the SCS POFDs would have been missed if we had only conducted a latitudinal study. These results help reveal the importance of assessing WMMCA performance for specific regions (e.g., regions with specific climatological characteristics, background surface types, etc.).

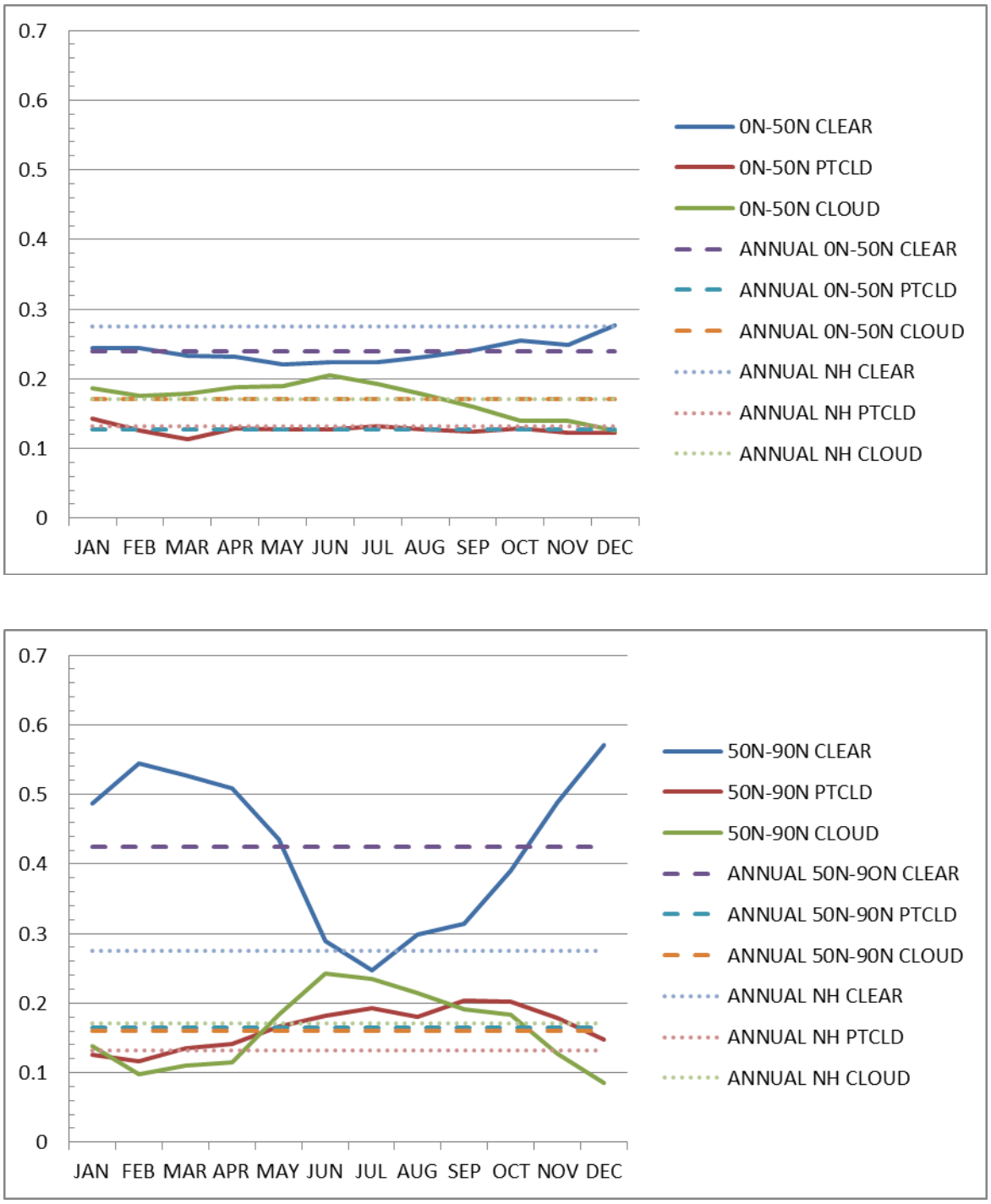


Figure 30. Probability of False Detection for each month and all three cloud categories. The top panel is for 0N–50N, while the bottom panel is for 50N–90N. This figure should be compared to Figure 31. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

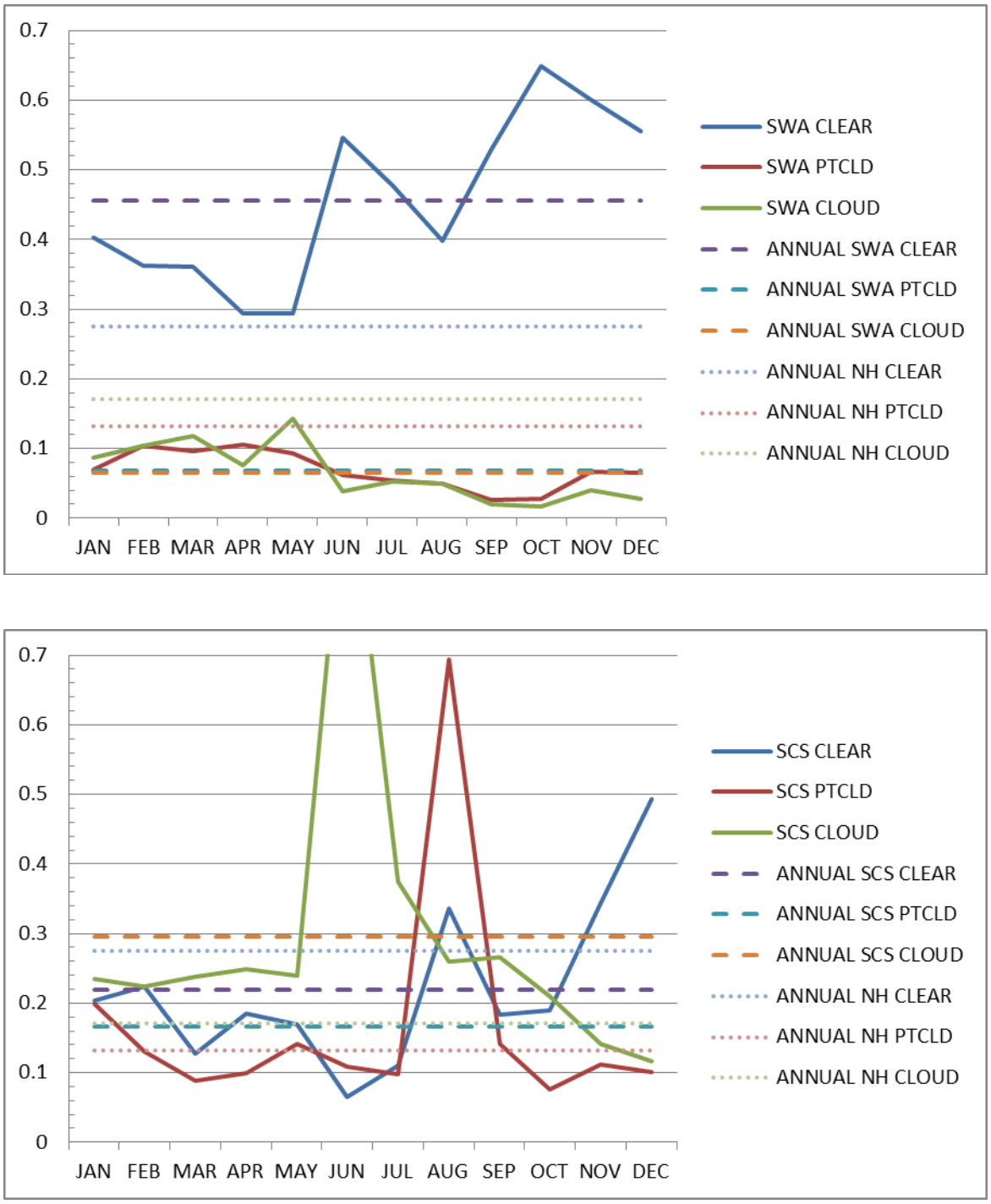


Figure 31. Probability of False Detection for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figure 30. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average. The POFD for Cloudy in SCS in June was 99%.

4. Bias

Figure 32 and Figure 33 show the bias results for each month and cloud category compiled by region. The annual Northern Hemisphere bias values were 1.09 for Clear, 0.95 for Cloudy and 0.86 for Partly Cloudy conditions. For the NH, Clear conditions always had a positive bias, Partly Cloudy conditions always negative bias and Cloudy conditions had both positive and negative bias depending on month. On a hemispheric scale, WWMCA appears to have only small bias in analyzing cloud categories. However, biases off by +/- 0.2 were seen in the high latitudes and in the smaller geographic regions. Bias at high latitudes were as low as 0.38 for fall Cloudy conditions and as high as 2.6 for Clear conditions during the same period. With the exception of the late spring and summer months, the bias for the high latitudes in all categories differed more than +/- 0.2 from the non-bias value of one. This bias at high latitudes needs to be addressed and adjusted for. Bias in the high latitudes is the only time where there may be the smallest discrimination and forecasting issues with the Partly Cloudy conditions versus any other cloud category with an annual average bias of one. This average bias value fails to capture the monthly trends in over and under biased events. Comparing bias values between the cloud categories reflects that where there is a tendency to over analyze Cloudy (Clear) conditions then there is also a tendency to under analyze Clear (Cloudy), though not to the same deviation from an unbiased analysis.

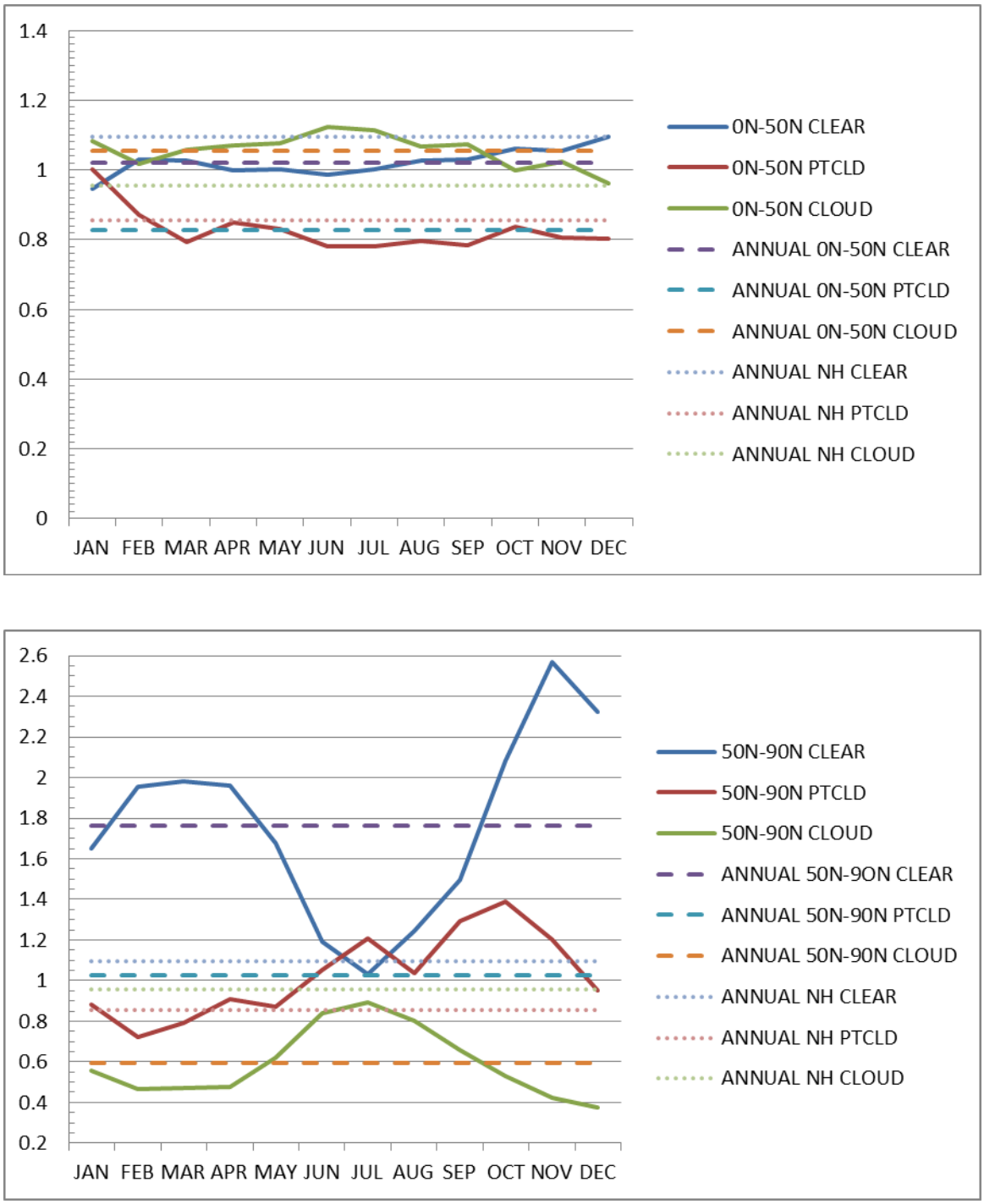


Figure 32. Bias results for each month and all three cloud categories. The top panel is for 0N–50N while the bottom panel is for 50N–90N. Note the y-axis variation to allow for larger range of bias at higher latitudes. This figure should be compared to Figure 33. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

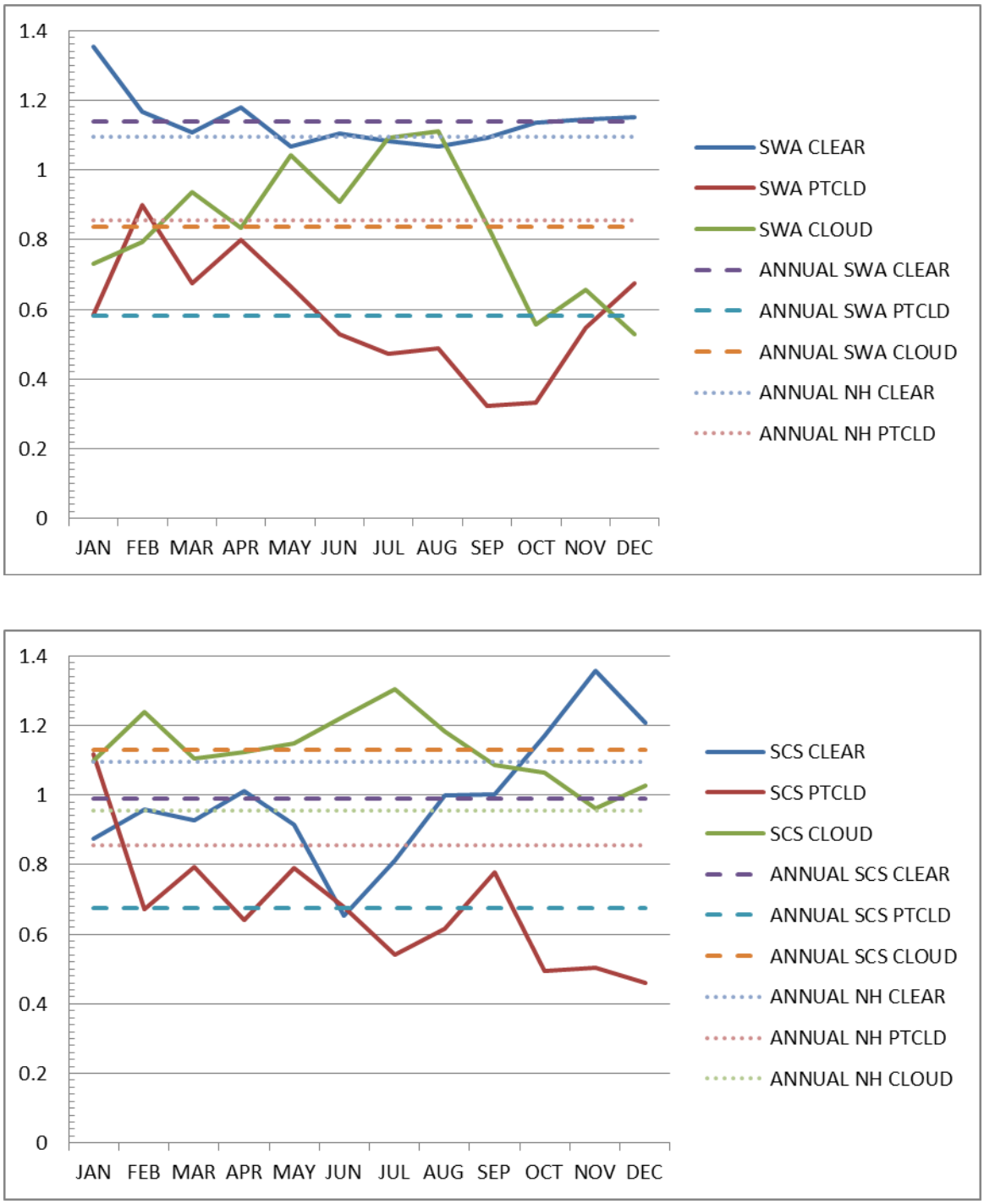


Figure 33. Bias results for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figure 32. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

5. Heidke Skill Score

Figure 35 and Figure 36 show the Heidke skill score for each month and cloud category by region. The annual Northern Hemisphere Heidke skill scores were 0.47 for Clear, 51% for Cloud, and 10% for Partly Cloudy, with monthly variations +/- 9% for Cloudy and Clear conditions, and +/- 2% for partly cloudy conditions. HSSs were best for cloudy conditions, followed by clear conditions, with partly cloudy conditions reporting the lowest skill. The skill at high latitudes is almost identical for both cloudy and clear categories and is on average 0.15–0.35 lower than the hemispheric average and worse than low latitudes. For latitudinal bands, HSS scores were always positive. There were negative HSSs for the two focus regions SWA and SCS over partly cloudy conditions. The largest seasonal variations in HSS for latitude bands were seen at the high latitudes. There is no set threshold of HSS that AFWA aims to achieve in its analyses. However HSS values over 50% might be considered a minimum threshold or benchmark that AFWA should aim to achieve for its near real time analyses, based on prior studies of HSS for short range forecasts issued by AFWA (e.g., Jarry 2005). Overall HSSs ranged between 0.48 and 0.62 between all regions except the high latitudes in the clear and cloudy categories indicating that there is good skill in the WWMCA forecast for these events.

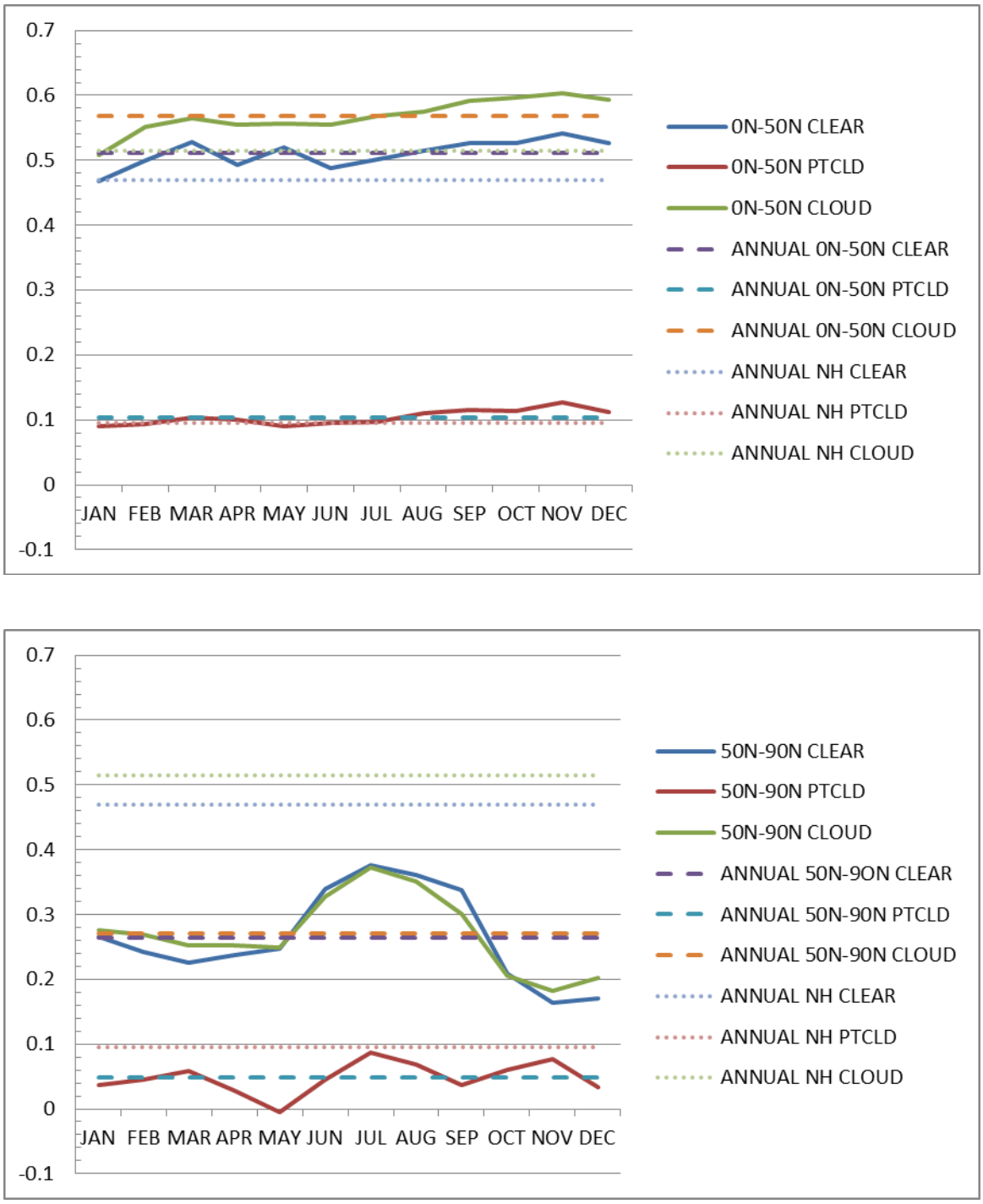


Figure 34. Heidke Skill Scores for each month and all three cloud categories. The top panel is for 0N–50N while the bottom panel is for 50N–90N. This figure should be compared to Figure 35. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

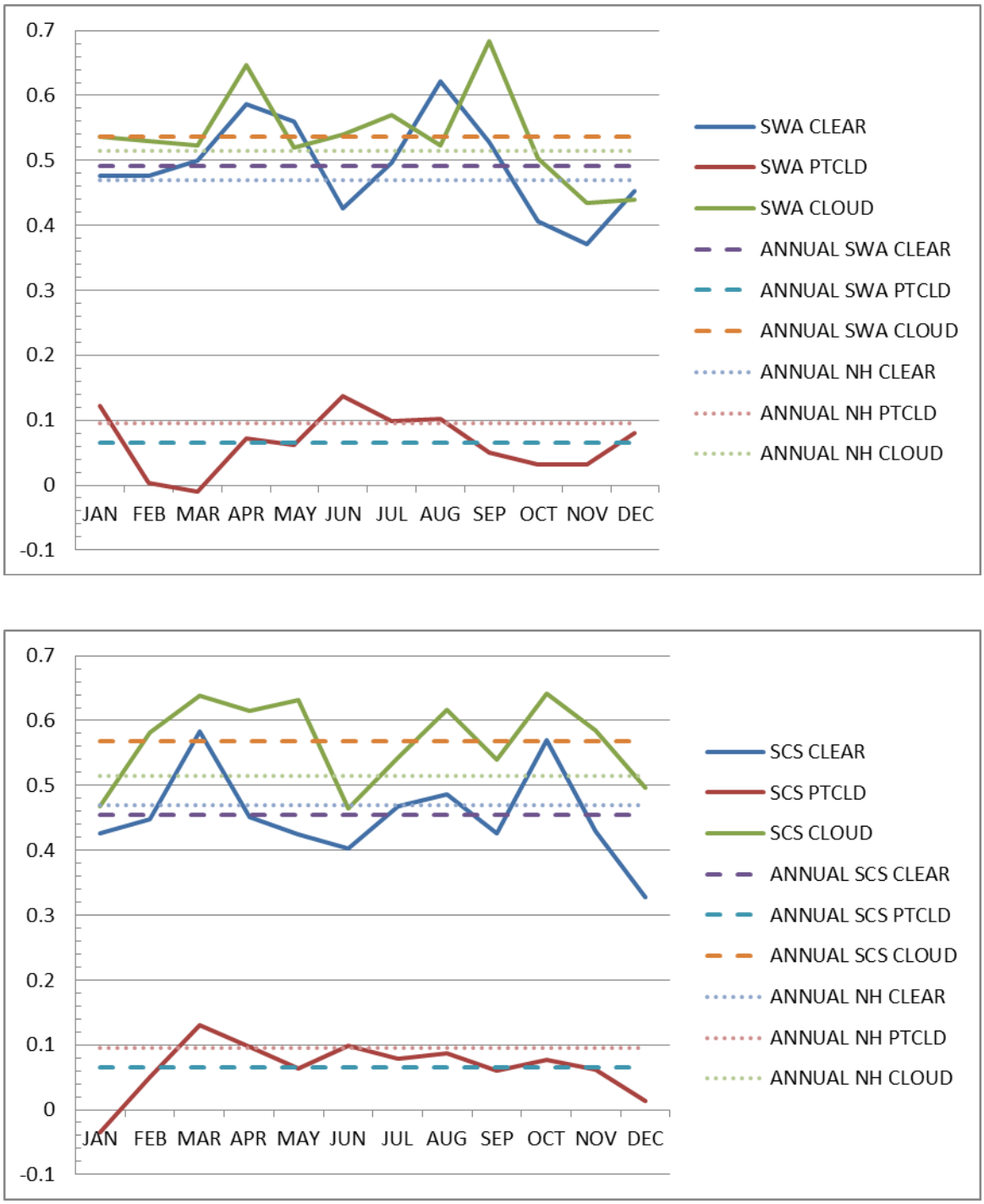


Figure 35. Heidke Skill Scores for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figure 34. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

6. Comparison of SWA Results to Prior Studies

Having latitudinal scale results is useful over hemispheric results however, results may not directly relate to a smaller scale interest area within the latitude band. Smaller scale studies were done on SWA and SCS to compare them to their latitude bands. Additionally, data from the SWA study can be compared to results from a similar study done by Cleary in 2012. Cleary studied the same region in 2010 over four months (Jan, Apr, Jul, Oct) and calculated similar performance metrics. POD values over SWA were within 5% between our study and Cleary for Clear conditions with slightly better POD for Cloudy conditions. Cleary determined that no bias existed during July for Cloudy and Clear conditions; however, our findings show that bias does exist during this time, but that it is at a minimum during the summer months. Additionally, overall bias results between our study and Cleary 2012 were greatly different (up to 40% different) for all cloud categories in these months. Cleary saw the best HSS for Clear conditions across Jan, Apr, Jul, and Oct while we saw the best HSS vary by month between Clear and Cloudy conditions. We also detected a greater range in the HSS scores with a 15–20% variance in HSS versus a 5–10% range between maximum and minimum scores. Both studies covered the exact same WWMCA box and months, and used CloudSat for validation. Because of the similarities in the base data set, it is most likely due to the differences in our collection and reduction methods and the use of 20 versus 30 for a cloud mask threshold that we see different results for the same performance metrics.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. SUMMARY, CONCLUSION AND RECOMMENDATIONS

A. SCOPE OF RESEARCH

A major goal of our study was to contribute to the improvement of cloud analyses and forecasts, and to thereby improve the planning and outcomes of DoD and IC operations. We conducted an analysis of WWMCA performance, or operational health, by comparing it to CloudSat as an independent data source for a number of locations and periods, using a range of performance metrics for all months of 2010. We aimed to overcome several of the shortcomings in this area of research that were identified in prior studies by considering a hemispheric, latitudinal, and focus region approach and including the calculation and comparison of 11 metrics. Additionally, we incorporated new approaches for data set reduction, including consideration of latitudinal variations in a WWMCA cell size and total data coverage ratio between the "truth" and analyzed data set. These changes and approaches were intended to better provide a well-rounded study which could be used to establish solid base-line performance metrics in various measures for use in improving WMMCA and for comparisons in and to future studies.

B. CONCLUSIONS

1. General Performance

WWMCA had an annual Northern Hemisphere hit proportions of 63% and miss proportion of 37%. WWMCA was within one cloud category in 24% of the occurrences and missed by two categories in 12% of the occurrences. The proportion correct for the Northern Hemisphere was best in March and September. Overall performance was better than the Northern Hemispheric average at the equator and worsened as latitude increased, with performance in mid-latitudes and high latitudes being worse than the Northern Hemisphere average. This could be due to a combination of factors, including latitudinal variations in the timeliness of data, data coverage and quantity, surface

backgrounds, seasonal effects, and predominant weather patterns and climate regimes. There were some large regional differences (e.g., between SWA and SCS) and large intraseasonal differences (e.g., between January and December) in performance, indicating that high spatial and temporal resolution is needed to fully characterize WMMCA.

2. Latency

Our approach to data time latency for WMMCA input data was twofold: assessments of performance based on adding older data to younger data, and on adding younger data to older data. Using both approaches revealed two significant jumps in WWMCA performance occurring at the 45 and 75 minute data ages. Data older than 75 minutes did little to improve the overall performance of the WWMCA, but not having any data for those regions would be detrimental and equally as bad, if not worse, than using older data. Due to the average temporal and spatial resolution of WWMCA data points, it may be useful to consider doing latitudinal time averaging above and below 50 degrees or at various latitude bands to create confidence levels by latitude, instead of looking at overall WWMCA pixel age average time to determine the confidence level or reliability of WWMCA. Product performance is best when input data is under 45 minutes old. At lower latitudes, most of the data is under 45 minutes old, which is probably a major reason why WMMCA higher performance tends to be higher in the lower latitudes. Knowing if data is younger and typically more reliable, or older and less reliable, for an individual focus region may be useful in determining the weight and confidence users should have in WMMCA for that region.

3. Cloud Categories

Verification metrics such as probability of detection, proportion correct, false alarm ratio, and probability of false detection varied between regions and seasons. While these metrics may have been higher for some regions and cloud categories than others, the HSS was used to determine best overall WWMCA

performance for cloud categories. The overall performance and Heidke skill scores demonstrated that WWMCA performs best in analyzing cloudy (>80%) conditions. Bias values tended to change between negative and positive as the seasons progressed. This could be due to adjustments made in the tuning algorithms to try to take into account changes in the surface background conditions. Table 8 summarizes the monthly ranges of the performance metrics for all regions, cloud categories, and performance metrics, including those not discussed in Chapter III. Additional results are available in Appendix C.

Table 8. Monthly ranges in performance metrics for the eight regions of study.

Area	Cloud Category	PC	POD	POFD	FAR	TS	BIAS	HSS
0N-90N	**Clear**	0.72-0.75	0.70-0.80	0.23-.034	0.29-0.33	0.53-0.58	1.01-1.17	0.43-0.50
	Pt Cld	0.74-0.77	0.20-0.24	.013-.014	0.72-0.77	0.12-0.15	0.79-0.98	0.07-0.12
	Cloudy	0.75-0.79	0.61-0.74	0.12-0.21	0.25-0.34	0.49-0.55	0.82-1.07	0.47-0.55
0N–50N	**Clear**	0.73-0.77	0.71-0.81	0.22-0.28	0.25-0.28	0.56-0.63	0.95-1.10	0.46-0.54
	Pt Cld	0.74-0.78	0.21-0.24	0.11-0.14	0.70-0.77	0.13-0.15	0.78-1.00	0.09-0.13
	Cloudy	0.78-0.82	0.71-0.77	0.13-0.21	0.28-0.35	0.51-0.58	0.96-1.12	0.50-0.60
50N–90N	**Clear**	0.53-0.70	0.63-0.84	0.25-0.57	0.39-0.67	0.26-0.45	1.02-2.57	0.16-0.38
	Pt Cld	0.70-0.77	0.16-0.29	0.11-0.20	0.76-0.82	0.09-0.15	0.72-1.39	-0.004-0.09
	Cloudy	0.54-0.69	0.31-0.57	0.09-0.24	0.18-0.33	0.29-0.47	0.37-0.89	0.18-0.37
0N–23.5N	**Clear**	0.73-0.77	0.71-0.82	0.20-0.31	0.21-.032	0.54-0.65	0.91-1.09	0.45-0.53
	Pt Cld	0.75-0.78	0.18-0.22	0.11-0.13	0.70-0.78	0.11-0.14	0.68-0.89	0.07-0.12
	Cloudy	0.79-0.84	0.73-0.82	0.12-0.21	0.25-0.39	0.50-.064	1.04-1.25	0.53-0.61
23.5N–35N	**Clear**	0.74-0.77	0.71-0.85	0.20-0.31	0.23-0.28	0.57-0.66	0.97-1.14	0.47-.052
	Pt Cld	0.73-0.78	0.21-0.28	0.12-0.17	0.67-.076	0.13-0.18	0.78-1.04	0.08-0.16
	Cloudy	0.77-0.84	0.64-0.78	0.08-0.19	0.22-.038	0.47-0.61	0.86-1.10	0.46-0.61
35N–50N	**Clear**	0.73-0.77	0.63-0.74	0.15-0.24	.022-0.46	0.43-0.58	0.85-1.17	0.41-0.53
	Pt Cld	0.71-0.78	0.25-0.32	0.13-0.18	0.72-0.76	0.14-0.17	0.93-1.19	0.08-0.15
	Cloudy	0.72-0.79	0.66-0.78	0.16-0.26	0.20-0.40	0.51-0.61	0.85-1.13	0.44-0.55
SWA	**Clear**	0.73-0.87	0.83-0.97	0.29-0.65	0.10-0.34	0.61-0.87	1.07-1.18	0.37-0.59
	Pt Cld	0.79-0.89	0.05-0.17	0.03-0.10	0.67-0.88	0.06-0.12	0.32-0.90	-0.009-0.13
	Cloudy	0.78-0.95	0.38-0.70	0.02-0.12	0.18-0.46	0.33-0.55	0.52-1.11	0.43-0.68
SCS	**Clear**	0.68-0.81	0.41-0.81	0.07-0.49	0.18-0.41	0.34-0.57	0.81-1.36	0.33-0.57
	Pt Cld	0.69-0.81	0.11-0.21	0.09-0.69	0.11-0.86	0.08-0.13	0.46-1.11	-0.04-0.1
	Cloudy	0.75-0.83	0.62-0.92	0.11-1	0.20-0.80	0.44-0.71	0.96-1.30	0.46-0.64

C. RECOMMENDATIONS FOR FUTURE RESEARCH

WWMCA is an ever-changing product, with the last major upgrade made in 2009. Additional upgrades have been made since 2011. A study of more recent data could use the 2010 study as a baseline to assess if the modifications are improving or degrading the performance of WWMCA.

We investigated WWMCA performance only for the Northern Hemisphere. It would be beneficial to replicate our study for the Southern Hemisphere. This would allow comparisons to be made to the UCAR study for the same period, as well as providing a more global assessment of WMMCA performance.

We focused on assessing WWMCA performance averaged over the whole Northern Hemisphere and latitudinal bands. WWMCA performance for more focused regions was only done for two relatively large regions, Southwest Asia and the South China Sea. We recommend that future research investigate performance in specific regions of interest to DoD and the IC. In addition to looking at specific AORs based on these interests, it would be useful to focus studies on areas of surface backgrounds and weather / climate conditions to better understand the behavior of WWMCA in more and less difficult challenging situations (situations with: (a) predominantly water, land, sand, snow, or ice surfaces; situations with well-organized weather systems, persistent clear skies, rapidly changing cloud conditions; etc.).

We did not conduct separate day and night assessments of WWMCA performance. We recommend that this be done because satellites sensors may change their preferred wavelength for cloud detection in periods of light versus periods of darkness. Results from such a study could lead to a change in the weight given during day and night hours to each input in the merging algorithm. Currently CloudSat is only providing data for the day time.

The current process for merging satellite information within and among satellite families only uses timeliness of data to determine data superiority. Instead of only using the age of the satellite data, it would be beneficial to consider both age and sensor capabilities of the data (e.g., slightly older polar orbiter time stamp versus newer geostationary time stamp, and multi-channel data versus single channel data). This is a more complex approach for determining data superiority, compared to the current use of pixel age alone, and could improve the accuracy of WWMCA. Since algorithms currently do not exist to do such weighing, we recommend a study be done to help establish these algorithms. Using the GDRs would be a good way to initiate such algorithms. A validation on the GDRs that are made by each family would provide critical information for modelers to determine how to modify the level 4 algorithms based on sensor performance, for use in fine tuning WWMCA output and improving the overall accuracy of CDFS II. By using CloudSat as truth, analyses could be done of the GDRs to determine the impacts of the time latency of each satellite sensor on WWMCA accuracy. This knowledge could then been incorporated into the decision algorithm CDFS II uses in combining the various satellite products to ensure the optimal set of observational data is used by WWMCA. Unfortunately, due to the large file size, higher frequency updates and limited AFWA storage capacity, the GDRs are not archived. The data needed for analyses of GDRs would need to be specifically requested prior to the creation of the GDRs. A temporary archive of GDRs was established for May–July 2012 in hopes of doing such a comparison study; however, CloudSat data was not made available during the same collection period by the start of our study.

There is no standardization for verification methods or metrics for WWMCA. It appears there is strong desire to learn more about the details of its performance in many subcategories. Additionally, CloudSat is nearing the end of its life, and may not always be around for comparison, and may not be the best source of truth data. We recommend comparisons be done between MODIS and CloudSat to determine if one, both, or neither of these systems should be

considered a good provider of data on actual cloud conditions. Finally, we recommend determining a standard averaging process for data timeliness and data coverage, if these are to be used as truth in the future for evaluation of WWMCA or other automated cloud forecasts and analyses.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. "CLOUD DEPICTION FORECAST SYSTEM (CDFS) II PROCESSING LEVELS"

This appendix describes the processes that occur throughout the four main levels of CDFSII. This appendix originally appeared in Cleary (2012) and has been modified only to adjust the formatting as well as provide an introduction.

> Satellite and conventional observations (surface observations and upper air soundings) undergo a four level process to be merged into a global cloud analysis (Figure 1). Level one is data calibration, level two classifies each pixel into cloudy or clear, level three applies cloud layering and typing, and level four consists of merging the separate analyses into one global analysis (HQ AFWA/DNXM 2011).

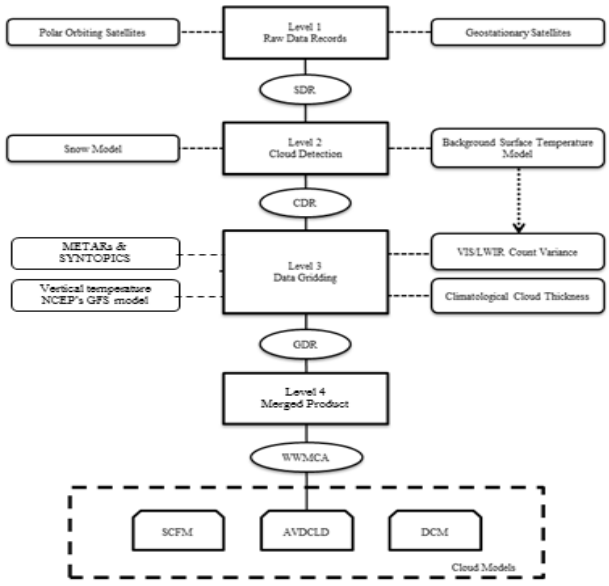


Figure 1. Illustration of the four processing levels within Cloud Depiction Forecast System II. Observations are received from meteorological satellites, conventional observations, and global analysis from various models. These observations are merged into one global cloud analysis that is used to initiate cloud forecast models. Shapes are defined as: rectangles are processes; rounded rectangles are inputs; ovals are products; and snipe same side corner rectangles are cloud models.

A. LEVEL 1

Level one processing consists of data ingestion and calibration. Satellite telemetry transmissions are received by AFWA's Satellite Data Handling System located at Offutt Air Force Base in Nebraska. Downlinked satellite data is encoded and must be decoded. The decoded data reveal physical parameters for radiance measurements received by the detectors on the sensor's focal plane array and are placed in the Sensor Data Records (SDR). Satellite imagers' focal plane array consist of detectors that represent the pixels, which measure the radiance received from reflected or emitted energy within Earth's atmosphere. CDFS II use the data from the satellite's visible and infrared channels. A calibration step for the infrared data converts the measured emitted energy into either radiance or brightness temperature. Brightness temperature is the temperature of an object if it was radiating as a black body. The brightness temperature is the parameter required by CDFS II to make the analysis. Reflectance values are measured from the satellite's visible channels are used directly in the algorithms (HQ AFWA/DNXM 2011).

B. LEVEL 2

Level two is where cloud detection occurs. Each sensor has its own tailored algorithms designed to optimize their instrument's ability to exploit measurements made in different channels in an attempt to distinguish cloud from clear scene. DMSP's Operational Line Scanner (OLS) sensor, has the highest spatial resolution, but only two broadband channels (one visible and one infrared), whereas, NOAA POES' Advanced Very-High Resolution Radiometer (AVHRR), have six narrowband channels (one visible, one near-infrared, and four infrared). DMSP's OLS cloud detection is accomplished by comparing a pixel's brightness temperature to a cloud-free referenced pixel's brightness temperature. If the pixel is determined to be cloud free, its brightness temperature is then used as the clear-scene brightness temperature for all other pixels in the frame. To determine a cloud-filled pixel, the observed brightness temperature of the said pixel is compared to the predicted clear-scene brightness temperature. The difference in magnitude of brightness temperature determines if the pixel is cloud-filled or cloudfree. A similar method is used for the visible channel when available. Threshold values are used to determine the cutoff between cloudy and clear pixel (HQ AFWA/DNMX 2011).

The algorithm for NOAA's AVHRR exploits the multispectral properties of the sensor to identify cloud or clear scene. The strength of the AVHRR is its six channels designs, which are listed in Table 2 with their respective wavelengths and typical use. Each channel differs in their sensitivity to reflectivity and emissivity properties of clouds and clear terrestrial surfaces. In addition to the sensors data, the algorithm uses clear-scene characterizations of the terrestrial background. The algorithm utilizes various techniques to include straight threshold type algorithms, inter-channel comparisons and spectral comparisons between the terrestrial surface and satellite data. A suite of twelve tests are used to characterize the different spectral characteristics of clouds and background surfaces to determine cloud-filled or clear scene. Each test is based on one or more specific spectral signatures that compare the radiance measurement of one or more channels, and fall into either cloud tests or background tests categories (HQ AFWA/DNXM 2011).

Table 1. The channels below are from the AVHRR/3 sensor. The AVHRR is a radiation-detection imager that can be used for remotely determining cloud cover and the surface temperature. Note that the term *surface* can mean the surface of the Earth, the upper surfaces of clouds, or the surface of a body of water. This scanning radiometer uses 6 detectors that collect different bands of radiation wavelengths as shown below. Measuring the same view, this array of diverse wavelengths, after processing, permits multi spectral analysis for more precisely defining hydrologic, oceanographic, and meteorological parameters. Comparison of data from two channels is often used to observe features or measure various environmental parameters. The three channels operating entirely within the infrared band are used to detect the heat radiation from and hence, the temperature of land, water, sea surfaces, and the clouds above them. Table based from NOAASIS (2011).

Channel number	Resolution at nadir (km)	Wavelength (μm)	Typical use
1	1.09	0.58-0.68	Daytime cloud and surface mapping
2	1.09	0.725-1.00	Land-water boundaries
3a	1.09	1.58-1.64	Snow and ice detection
3b	1.09	3.55-3.93	Night cloud mapping, sea surface temperature
4	1.09	10.30-11.30	Night cloud mapping, sea surface temperature
5	1.09	11.50-12.50	Sea surface temperature

There are nine cloud tests and three background tests, which are summarized in Table 3. Different tests are used to identify clouds under different conditions. The low clouds and fog test for solar-illuminated data is used to identify water droplets based low-level cloud when the scene is illuminated by sunlight, and the non-

illuminated test is used during nighttime. Since no one test will identify all the clouds in a scene, the cloud tests must be used in combination to accurately identify all cloud-filled pixels. The background tests are unique to the AVHRR algorithm, which exploits the multispectral characteristics of AVHRR data, to identify snow and ice, desert and sun glint backgrounds. These tests are essential because clouds and surface features often exhibit similar spectral signatures in the visible spectrum, however; a positive result from these tests does not automatic mean a cloudfree pixel. These tests identify suspected visible data; the infrared cloud tests must still be applied to determine a cloud-filled pixel (HQ AFWA/DNXM 2011).

Table 2. Cloud analysis test for the NOAA AVHRR level 2 algorithm. Table from HQ AFWA/DNXM (2011).

Test Type	Test Name	Major Identifying Class	Visible Data	IR Data
Cloud	Low cloud and fog test for solar-illumination data	Solar-illumination liquid water cloud		X
	Precipitation cloud test	Cumulonimbus	X	X
	Thin cirrus cloud test for solar-illumination data	Solar-illumination ice cloud	X	X
	Visible brightness ratio	Solar-illuminated liquid water cloud	X	
	Single channel visible brightness test	Solar-illuminated liquid water cloud	X	
	Cold cloud test	Mid- to high-level optically thick water and ice cloud		X
	Cirrus cloud test	High-level ice cloud		X
	Fog, low stratus test for non-solar-illuminated data	Non-solar-illuminated liquid water cloud		X
	Thin cirrus cloud for non-solar-illuminated data	Non-solar-illuminated ice cloud		X
Background	Sun glint background test	Water surfaces exhibiting specular reflection	X	X
	Desert background test	Highly reflective non-vegetated land surfaces	X	X
	Snow/ice cover background test	Highly reflective snow or ice covered land and water surfaces	X	X

There are three algorithms used to detection clouds from geostationary satellite data. Geostationary satellites have a high temporal resolution but the spatial resolution is degraded due to its altitude (~36,000 km). The first algorithm takes advantage of the high temporal resolution to identify cloud-filled pixels by testing for rapid changes in brightness temperature and reflectance values in pixels representing the same geolocation. The pixels that exhibit changes in radiance values greater than the amount expected for clear scene from frame to frame are identified as cloud-filled. The second algorithm is a dynamic threshold algorithm that identifies cloud with similar characteristics. Cloud-filled pixels identified through the temporal difference algorithm are processed by the dynamic threshold algorithm. The dynamic threshold algorithm

identifies maximum and minimum brightness temperatures or reflectance within a grid cell, which are used to define threshold values for cloud-filled and cloud-free pixels remaining within the grid cell. The third algorithm uses a series of spectral discrimination tests similar to the OLS and AVHRR spectral tests. Not all geostationary satellite data are the same, so a different set of tests may need to be ran for each satellite system. For instance, METOSAT platforms have different spectral channels than GOES (HQ AFWA/DNXM 2011). Tables 4 and 5 summarize the spectral channels for METEOSAT and GOES, respectively. The resultant dataset, generated for each satellite data source, is called Cloud Data Records (CDR).

Table 3. Spectral channels and bandwidth for METEOSAT satellites. Table from EUMETSAT (2011).

Channel	Wavelength (µm)	Spatial Resolution at nadir (km)	Remarks
VIS 0.6	0.56-0.71	3	Similar to AVHRR
VIS 0.8	0.74-0.88	3	Similar to AVHRR
IR 1.6	1.50-1.78	3	Similar to AVHRR
IR 3.9	3.48-4.36	3	Similar to AVHRR
IR 8.7	8.30-9.10	3	New
IR 10.8	9.80-11.80	3	Similar to AVHRR
IR 12.0	11.00-13.00	3	Similar to AVHRR
WV 6.2	5.35-7.15	3	Water vapor channel
WV 7.3	6.85-7.85	3	Water vapor channel
IR 9.7	9.38-9.94	3	Ozone absorption channel as on HIRS
IR 13.4	12.40-14.40	3	CO2 absorption channel as on GOES-VAS sounder
HRV	0.5-0.9	1	High Resolution Visible (HRV); Broadband visible channel

Table 4. GOES imager channels. Table from GOES Imager Channel Notation (2011).

Channel number	Resolution at nadir or IFOV* (km)	Wavelength (µm)	Remarks
1	1	0.55-0.75	Visible
2	4	3.80-4.00	Shortwave Infrared
3	8	6.50-7.00	Moisture
4	4	10.20-11.20	Infrared 1
5	4	11.50-12.50	Infrared 2

* Instantaneous Field of View

C. LEVEL 3

Level three is where the satellite pixels are gridded onto AFWA's standard Polar-stereographic grid at "16th mesh" with a horizontal resolution of 24 km (true at 60° latitude). Pixels are assembled into the 16th mesh grid cells by computing the coordinates that correspond to the latitude and longitude of each pixel (Hoke et al., 1981 Rev. March 1985). A detailed description of the Polar-stereographic grid is provided in Map Projections and Grid System for Meteorological Applications, *AFGWC Technical Notes 79/003* (Hoke et al., 1981 Rev. March 1985). The cloud layers are identified through the Long Wave Infrared (LWIR) brightness temperature data contained within each grid cell. A clustering algorithm clusters pixels of similar brightness characteristics to identify potential layer separations. Statistical procedures are applied to the grid cell to limit the identified layers to four. Once the layers are identified, cloud top temperatures are compared against vertical temperature information from the Nation Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) model to assign a cloud top height. The cloud top height and temperature information is used with the visible/LWIR-count variance, from the background surface temperature model employed in level two, to assign each layer to one of nine different cloud types listed in Table 6. Along with each derived cloud type is a climatological cloud thickness that is subtracted from the cloud top height to determine the cloud base height (HQ AFWA/DNXM 2011). At the end of level three, each satellite family (e.g., DMSP, NOAA, Geostationary) have a common gridded cloud mask that consists of cloud fraction up to four layers. The cloud masking also includes cloud type, and cloud top/base heights. These datasets are called Gridded Data Records (GDR).

Table 5. WWMCA default cloud thickness according to height. Cloud thickness is based on climatology. Table from HQ AFWA/DNXM (2011).

Type Code	Cloud Type	Thickness (m)
1	Cumulonimbus (Cb)	6300
2	Strata (St)	600
3	Stratocumulus (Sc)	1800
4	Cumulus (Cu)	1200
5	Altostratus (As)	1200
6	Nimbostratus (Ns)	2100
7	Altocumulus (Ac)	1800
8	Cirrostratus (Cs)	1800
9	Cirrus (Ci)	900

Level three processing also includes hourly global surface and upper air based data, METARS or SYNTOPIC type formats, which contain fractional cloud coverage and cloud base heights from the World Meteorological Organization. These conventional observations are combined with the satellite data to determine the cloud mask, cloud type, and cloud top/base heights.

D. LEVEL 4

Level four is where the satellite family GDRs and conventional surface observations are merged into a single global analysis of cloud cover information. One problem that arises in level four is that the independent gridded analyses have different valid times because the satellites input their data into CDFS II at different times. Each independent gridded analysis has strengths and weaknesses. For example, the polar-orbiters (DMSP and NOAA satellites) derived analyses have greater accuracy from the spatial resolution (polar satellites are in a lower orbit, ~800 km); however, the temporal resolution is course, usually passing over a particular region one or two times a day. Geostationary satellites analyses have a finer temporal resolution, every 30 minutes, but spatial resolution, or instantaneous field of view (IFOV), varies from 1 to 8 km depending on the channel. See the resolution at nadir column in Table 12 for each channel's IFOV. The timeliness and accuracy of the observations is a major concern when merging the data into a one global analysis (HQ AFWA/DNXM 2011).

Integration of total cloud amounts supersedes integration of layered quantities since total cloud fraction estimates are more accurate than individual layer fraction due to the sample size of total cloud amount is far greater than the layered cloud amounts (HQ AFWA/DNXM 2011). Bartlett (2009) explains how the total cloud amount analysis works:

> Total cloud fraction is then set to either 100 or 0 percent, respectively. If neither analysis is completely cloud-filled or completely cloud-free, then the error for each analysis is estimated. The estimated errors for the analyses are compared to one another to see if the most recent analysis also has the lowest estimated error. Optimum interpolation (OI) occurs when one analysis cannot be chosen as the most accurate. OI maintains a blended estimate of total cloud fraction from multiple input analyses. Weighting functions for the OI are based on the estimated analysis errors which are computed for each individual analysis. Analysis errors are defined as an initial analysis error plus an additional error growth function which grows linearly with time. The error growth function is a tunable parameter that analysts can adjust to correct for inconsistencies.

When total cloud amount is completed the other cloud parameters are merged. Rules applied, to determine which analysis is superior, in the layered analysis are similar to the total cloud amount; however, if multiple timely analyses have 100% cloud cover or it is determined that an OI technique is necessary. The integration of layered cloud amounts undergoes a more extensive algorithm. Most likely the individual analyses will have varying vertical distributions of cloud and cloud type due to the differences in sensor characteristics for each satellite family. The more complex algorithm determines which analyses is the most accurate and designates that analyses as the master template for which all other timely analyses are merged on. This process impacts discrete values such as the number of cloud layers and cloud types because these when integrated they will assume the values of the master template. The OI procedure is used for varying layered cloud faction and cloud top temperature. The OI process combines layers that closely match in cloud top temperatures and determines the layered cloud fraction. Special cloud algorithms have been designed for certain satellite sensors to enhance detection of low level stratus and cirrus. These special-case clouds are verified against the integrated analysis to be certain that the analysis is

accurate, and are effective in showing the persistence of the observations in the subsequent integration analysis (HQ AFWA/DNXM 2011).

All the output variables are placed in a GriB file (a collection of individual self-containing records, and the individual records themselves can stand alone as meaningful data) and is published as the Worldwide Merge Cloud Analysis (WWMCA).

APPENDIX B. MATLAB CODE EXAMPLES FOR DATA PROCESSING

This appendix provides examples for code used to process the data via MATLAB. Due to the length of the programs not all codes are written in their entirety. Please contact Dr. Tom Murphree in the Meteorology department at NPS for full copies of the code at murphree@nps.edu or (com) 831-656-2723.

A. MATLAB STEP 1 DATA FILE FORMAT QUALTIY CONTROL

```
% file = step1_data_fileformat_QC_16Nov.m
clear all
close all
clc
dir_ORIG = '\\comfort\cmlecomp$\Desktop\MATLAB_CURRENT\ORIGINAL_CSV\'
dir_QC   = '\\comfort\cmlecomp$\Desktop\MATLAB_CURRENT\QC_CSV\'
file_to_QC = 'CloudSat_wwmca_phase3_12_2010.csv'
file_after_QC = [file_to_QC(1:end-4) '_QCd.csv']

% make file names with directories
input_file = [dir_ORIG file_to_QC]
output_file = [dir_QC file_after_QC]

% ----------------------------------
% Step 1:  See if file has bad lines
% ----------------------------------
% Open Input File
fid = fopen(input_file,'r')

header_line = fgetl(fid)
icount = 0;
for i = 1:5e6
    line = fgetl(fid);
    if line == -1
        disp(['line number = ' int2str(i)])
        break
    end
    index = findstr(line,',');

    N = 23;   % for CloudSat WWMCA dataset from Bruce Ford
    if strcmp(line(11:11),'z') & length(index) ~= N
        disp(['line i = ' int2str(i) ' has N = '
int2str(length(index)) ' entries'])
        icount = icount+1;
    end
end
icount
fclose(fid);
```

```
% ------------------------------------------
% Step 2:  Write QC'd file to omit bad lines
% ------------------------------------------
 %  Write the output file without the Bad Lines
if icount > 0

    fid = fopen(input_file,'r')  % read only -- cannot write to it
    % open output file with write option
    fid_out = fopen(output_file,'w')

    % read header line
    header_line = fgetl(fid);
    % write header line to Output file
    fprintf(fid_out,'%s\n',header_line);

    icount = 0;
    for i = 1:5e6
        line = fgetl(fid);
        if line == -1
            disp(['line number = ' int2str(i)])
            break
        end
        index = findstr(line,',');

        N = 23;   % for CloudSat WWMCA dataset from Bruce Ford
        if strcmp(line(11:11),'z') & length(index) ~= N
            disp(['line i = ' int2str(i) ' has N = '
int2str(length(index)) ' entries'])
            icount = icount+1;
        else
            fprintf(fid_out,'%s\n',line);  % write to output file
        end
    end
    icount
    fclose(fid);
    fclose(fid_out);
else
    % if all the lines were Good, we will copy the file instead of
    % rewritting as above
    [copy_success,copy_message] = copyfile(input_file,output_file,'f')
end
```

B. MATLAB STEP 2 DECODE QUALTIY CONTROLLED DATA

```
% file = step2_DECODE_QC_CSV_16Nov12.m
clear all
close all
clc
dir_QC   = '\\comfort\cmlecomp$\Desktop\MATLAB_CURRENT\QC_CSV\'
file_after_QC = 'CloudSat_wwmca_phase3_12_2010_QCd.csv'  %USER CHANGE
HERE
input_file = [dir_QC file_after_QC]
dir_MAT = '\\comfort\cmlecomp$\Desktop\MATLAB_CURRENT\MAT_PHASE4\'
matfile = [dir_MAT file_after_QC(23:end-8) '.mat']
```

```
fid = fopen(input_file,'r')
header_line = fgetl(fid);

for i = 1:5e6
    line = fgetl(fid);
    if line == -1
        disp(['line number = ' int2str(i)])
        break
    end
    index = findstr(line,',');

    % extract YYYYMMDDHH as string
    YYYYMMDDHH_str(i,:) = line(1:10);
    % column 2
    AVG_CSLAT(i,1) = str2num(line(index(1)+1:index(2)-1));
    % column 3
    WWMCA_I(i,1) = str2num(line(index(2)+1:index(3)-1));
    % column 4
    WWMCA_J(i,1) = str2num(line(index(3)+1:index(4)-1));
    % column 5
    WWMCA_BOX(i,1) = str2num(line(index(4)+1:index(5)-1));
    % column 6
    BOX_LENGTH(i,1) = str2num(line(index(5)+1:index(6)-1));
    % column 7
    BOX_AREA(i,1) = str2num(line(index(6)+1:index(7)-1));
    % column 8
    BOX_COVERAGE(i,1) = str2num(line(index(7)+1:index(8)-1));
    % column 9 ... character string D = day, N = night
    str = line(index(8)+1:index(9)-1);
    if strcmp(str,'N')
        DAY_NIGHT_FLAG(i,1) = -1;  %  Night = -1
    elseif strcmp(str,'D')
        DAY_NIGHT_FLAG(i,1) = 1;  %  Day = 1
    end

    % column 10 ... character string 50 degree flag A = above, B =
below
    str = line(index(9)+1:index(10)-1);
    if strcmp(str,'A')
        DEG_MARKER(i,1) = 1;  %  Above = 1
    elseif strcmp(str,'B')
        DEG_MARKER(i,1) = -1;  %  Below = -1
    end

    % column 11
    N_CSP_TOTAL(i,1) = str2num(line(index(10)+1:index(11)-1));
   % column 12
    N_CSP_CLOUD(i,1) = str2num(line(index(11)+1:index(12)-1));
    % column 13
    N_CSP_PC(i,1) = str2num(line(index(12)+1:index(13)-1));
    % column 14
    N_CSP_NO_CLOUD(i,1) = str2num(line(index(13)+1:index(14)-1));
    % column 15
```

```
    SUM_CSP(i,1) = str2num(line(index(14)+1:index(15)-1));
    % column 16
    CS_20_TCC(i,1) = str2num(line(index(15)+1:index(16)-1));
    % column 17
    CS_20_CCB(i,1) = str2num(line(index(16)+1:index(17)-1));
    % column 18
    CS_30_TCC(i,1) = str2num(line(index(17)+1:index(18)-1));
    % column 19
    CS_30_CCB(i,1) = str2num(line(index(18)+1:index(19)-1));
    % column 20
    WWMCA_LAT(i,1) = str2num(line(index(19)+1:index(20)-1));
    % column 21
    WWMCA_LONG(i,1) = str2num(line(index(20)+1:index(21)-1));
    % column 22
    WWMCA_CCB(i,1) = str2num(line(index(21)+1:index(22)-1));
    % column 23
    WWMCA_TCC(i,1) = str2num(line(index(22)+1:index(23)-1));
    % column 24
    WWMCA_PIXEL_AGE(i,1) = str2num(line(index(23)+1:end));

end
fclose(fid);

clear ans i index line str fid
clear dir_QC file_after_QC input_file dir_MAT

YYYY = str2num(YYYYMMDDHH_str(:,1:4));
MM = str2num(YYYYMMDDHH_str(:,5:6));
DD = str2num(YYYYMMDDHH_str(:,7:8));
HH = str2num(YYYYMMDDHH_str(:,9:10));

whos

eval(['save  ' matfile])
%
```

C. MATLAB STEP 3A FUNCTION TO CALCULATE

Note that the steps under the " % NORTHERN HEMISPHERE TOTAL" were repeated in this function to account for all regions.

```
function [SUM_TOTAL] =
step3_FUNCTION_to_CALCULATE_REGION_16Jan(matfile,index1,index2,index3,i
ndex4,index5,index6,index7,index8)
% Chandra LeCompte thesis calculations
disp(' ************ Entering Function ******************')
% Load MAT File -----------------------------------------
feval('load', matfile)
% NORTHERN HEMISPHERE TOTAL  ----------------------------------------
TRUTH_20 = CS_20_CCB (index1);
N_TRUTH_20 = length(TRUTH_20)
FCST = WWMCA_CCB (index1);
```

```
N_FCST = length(FCST)
OVERALL_20_HIT = sum(TRUTH_20 == 0 & FCST == 0 |TRUTH_20 == 1 & FCST ==
1 |TRUTH_20 == 2 & FCST ==2)% number of times the cloud category was
properly forecasted and observed
OVERALL_20_1_CAT_MISS = sum(TRUTH_20 == 1 & FCST ~=1 | TRUTH_20 ~= 1 &
FCST == 1)% number of times the forecast was off from the observed by
one category of clouds
OVERALL_20_2_CAT_MISS = sum(TRUTH_20 == 2 & FCST == 0 | TRUTH_20 == 0 &
FCST == 2)% number of times the forecast was off from the observed by
two categories of clouds , (ie forecasted no cloud when cloud, or cloud
when no cloud)

NC_20_HIT = sum(TRUTH_20 == 0 & FCST == 0) % number of accurately
forecasted no cloud
NC_20_MISS = sum(TRUTH_20 == 0 & (FCST == 1 | FCST == 2)) % forecasted
cloud but truth is No Cloud
NC_20_FA = sum(TRUTH_20 ~= 0 & FCST == 0)% forecasted no cloud but
cloud occured
NC_20_CR = sum((TRUTH_20 == 1 |TRUTH_20 == 2) & (FCST == 1 | FCST ==
2)) % fcst and truth are Cloud, properly forecasted cloud to exist
clear XX_matrix POD PROP_COR TS BIAS FAR HSS

XX_matrix = [NC_20_HIT, NC_20_MISS, NC_20_FA, NC_20_CR]
TOTAL = (XX_matrix(1)+XX_matrix(2)+ XX_matrix(3) + XX_matrix(4))
POD = XX_matrix(1)/(XX_matrix(1) + XX_matrix(2))  %will need to repeat
this for all 8 categories with proper H,M,CR,FA
PROP_COR = (XX_matrix(1) + XX_matrix(4))/(XX_matrix(1) + XX_matrix(2) +
XX_matrix(3) + XX_matrix(4))%will need to repeat this for all 8
categories with proper H,M,CR,FA
TS = (XX_matrix(1))/(XX_matrix(1)+ XX_matrix(2)+ XX_matrix(3))
BIAS = (XX_matrix(1)+XX_matrix(3))/(XX_matrix(1)+ XX_matrix(2))
FAR =
(XX_matrix(3))/(XX_matrix(1)+XX_matrix(2)+XX_matrix(3)+XX_matrix(4))%wi
ll need to repeat this for all 8 categories with proper H,M,CR,FA
HSS = (2*((XX_matrix(1)*XX_matrix(4))-
(XX_matrix(2)*XX_matrix(3))))/((XX_matrix(1)+XX_matrix(2))*(XX_matrix(2
) + XX_matrix(4))+((XX_matrix(1)+XX_matrix(3))*(XX_matrix(3) +
XX_matrix(4)))) %will need to repeat this for all 8 categories with
proper H,M,CR,FA
RESULT_NC_20 = [NC_20_HIT NC_20_MISS NC_20_FA NC_20_CR POD PROP_COR TS
BIAS FAR HSS]

PC_20_HIT = sum(TRUTH_20 == 1 & FCST == 1) % number of accurately
forecasted partly cloud
PC_20_MISS = sum(TRUTH_20 == 1 & (FCST == 0 | FCST == 2)) % forecasted
cloud or no cloud but truth is partly Cloud
PC_20_FA = sum(TRUTH_20 ~= 1 & FCST == 1)% forecasted partly cloud but
cloud or no cloud occured
PC_20_CR = sum((TRUTH_20 == 0 |TRUTH_20 == 2) & (FCST == 0 | FCST ==
2)) % fcst and truth are Cloud or no cloud, properly forecasted partly
cloud would not happen
clear XX_matrix POD PROP_COR TS BIAS FAR HSS

XX_matrix = [PC_20_HIT, PC_20_MISS, PC_20_FA, PC_20_CR]
TOTAL = (XX_matrix(1)+XX_matrix(2)+ XX_matrix(3) + XX_matrix(4))
```

```
POD = XX_matrix(1)/(XX_matrix(1) + XX_matrix(2))  %will need to repeat
this for all 8 categories with proper H,M,CR,FA
PROP_COR = (XX_matrix(1) + XX_matrix(4))/(XX_matrix(1) + XX_matrix(2) +
XX_matrix(3) + XX_matrix(4))%will need to repeat this for all 8
categories with proper H,M,CR,FA
TS = (XX_matrix(1))/(XX_matrix(1)+ XX_matrix(2)+ XX_matrix(3))
BIAS = (XX_matrix(1)+XX_matrix(3))/(XX_matrix(1)+ XX_matrix(2))
FAR =
(XX_matrix(3))/(XX_matrix(1)+XX_matrix(2)+XX_matrix(3)+XX_matrix(4))%wi
ll need to repeat this for all 8 categories with proper H,M,CR,FA
HSS = (2*((XX_matrix(1)*XX_matrix(4))-
(XX_matrix(2)*XX_matrix(3))))/((XX_matrix(1)+XX_matrix(2))*(XX_matrix(2
) + XX_matrix(4))+((XX_matrix(1)+XX_matrix(3))*(XX_matrix(3) +
XX_matrix(4)))) %will need to repeat this for all 8 categories with
proper H,M,CR,FA
RESULT_PC_20 = [PC_20_HIT PC_20_MISS PC_20_FA PC_20_CR POD PROP_COR TS
BIAS FAR HSS]

CLOUD_20_HIT = sum(TRUTH_20 == 2 & FCST == 2)% number of accurately
forecasted cloud
CLOUD_20_MISS = sum(TRUTH_20 == 2 & (FCST == 1 | FCST == 0)) %
forecasted no cloud but truth is Cloud
CLOUD_20_FA = sum(TRUTH_20 ~= 2 & FCST == 2)% forecasted cloud but no
cloud occured
CLOUD_20_CR = sum((TRUTH_20 == 1 |TRUTH_20 == 0) & (FCST == 1 | FCST ==
0)) % fcst and truth are no cloud, properly forecasted no cloud to
exist

clear XX_matrix POD PROP_COR TS BIAS FAR HSS
XX_matrix = [CLOUD_20_HIT, CLOUD_20_MISS, CLOUD_20_FA, CLOUD_20_CR]
TOTAL = (XX_matrix(1)+XX_matrix(2)+ XX_matrix(3) + XX_matrix(4))
POD = XX_matrix(1)/(XX_matrix(1) + XX_matrix(2))  %will need to repeat
this for all 8 categories with proper H,M,CR,FA
PROP_COR = (XX_matrix(1) + XX_matrix(4))/(XX_matrix(1) + XX_matrix(2) +
XX_matrix(3) + XX_matrix(4))%will need to repeat this for all 8
categories with proper H,M,CR,FA
TS = (XX_matrix(1))/(XX_matrix(1)+ XX_matrix(2)+ XX_matrix(3))
BIAS = (XX_matrix(1)+XX_matrix(3))/(XX_matrix(1)+ XX_matrix(2))
FAR =
(XX_matrix(3))/(XX_matrix(1)+XX_matrix(2)+XX_matrix(3)+XX_matrix(4))%wi
ll need to repeat this for all 8 categories with proper H,M,CR,FA
HSS = (2*((XX_matrix(1)*XX_matrix(4))-
(XX_matrix(2)*XX_matrix(3))))/((XX_matrix(1)+XX_matrix(2))*(XX_matrix(2
) + XX_matrix(4))+((XX_matrix(1)+XX_matrix(3))*(XX_matrix(3) +
XX_matrix(4)))) %will need to repeat this for all 8 categories with
proper H,M,CR,FA
RESULT_CLOUD_20 = [CLOUD_20_HIT CLOUD_20_MISS CLOUD_20_FA CLOUD_20_CR
POD PROP_COR TS BIAS FAR HSS]

TOTAL_RESULT_20 = [ N_FCST OVERALL_20_HIT OVERALL_20_1_CAT_MISS
OVERALL_20_2_CAT_MISS RESULT_NC_20 RESULT_PC_20 RESULT_CLOUD_20 ]

disp(' ************ Exiting Function ******************')
disp('   ')
% ---------- end of function ------------------------
```

D. MATLAB STEP 3B MAIN PROGRAM TO CALCULATE

Note that the “column” calculations were repeated in this function to account for all of the time block periods.

```
% file = step3_MAIN_PTC_REGION_16Jan.m
% Purpose:  MATRIX with greater than 6% box coverage, 20 threshold,
reduced time binning and all 3 displays, include region WWMCA box 22
and 12
% ---------------------------------------------------------------
clear all
close all
clc
format bank
dir_MAT = '\\comfort\cmlecomp$\Desktop\MATLAB_CURRENT\MAT_PHASE4\'
matfile = [dir_MAT '01_2010.mat'] % USER Change --------------
dir_OUTPUT = '\\comfort\cmlecomp$\Desktop\MATLAB_CURRENT\MATRIX\'
output_MATFILE = [dir_OUTPUT '01_2010_REGIONS_COR2.mat'] % USER Change
% load Phase4 MAT (input MAT)
feval('load', matfile)
% -------------------- -----------------------------
%  Column 1 240 <= AGE
% -------------------------------------------------
ICOLUMN = 1
index1 = (WWMCA_LAT<=90 & 240<=WWMCA_PIXEL_AGE & 6<=BOX_COVERAGE); %
this should give all points in the data set
index2 = (WWMCA_LAT<50 & 240<=WWMCA_PIXEL_AGE & 6<=BOX_COVERAGE); %
all points in the data set below 50 degrees
index3 = (WWMCA_LAT <23.5 & 240<=WWMCA_PIXEL_AGE & 6<=BOX_COVERAGE);
%tropics
index4 = ( 23.5<= WWMCA_LAT & WWMCA_LAT<35 & 240<=WWMCA_PIXEL_AGE &
6<=BOX_COVERAGE); %  all points in the data set for subtropics
index5 = ( 35<= WWMCA_LAT & WWMCA_LAT<50 & 240<=WWMCA_PIXEL_AGE &
6<=BOX_COVERAGE); %  all points in the data set for midlat
index6 = ( 50<= WWMCA_LAT & 240<=WWMCA_PIXEL_AGE & 6<=BOX_COVERAGE); %
all points in the data set for polar
index7 = (WWMCA_BOX == 22 & 240<=WWMCA_PIXEL_AGE & 6<=BOX_COVERAGE);
index8 = (WWMCA_BOX == 12 & 240<=WWMCA_PIXEL_AGE & 6<=BOX_COVERAGE);

[SUM_TOTAL] =
step3_FUNCTION_to_CALCULATE_REGION_16Jan(matfile,index1,index2,index3,i
ndex4,index5,index6,index7,index8); % USER CHANGE ALL LOCATIONS IF ADD
MORE INDEX OR CHANGE NAME OF THE FUNCTION

FINAL_MATRIX(:,ICOLUMN) = SUM_TOTAL'
clear SUM_TOTAL index*
eval(['save ' output_MATFILE ' FINAL_MATRIX '])
```

E. MATLAB STEP 4 WRITING TO EXCEL

```
% file = step4_write_to_EXCEL_20Dec12.m
% Purpose:
% ------------------------------------------------------------
clear all
close all
clc
format bank

dir_MAT = '\\comfort\cmlecomp$\Desktop\MATLAB_CURRENT\MATRIX\'
%matfile1 = [dir_MAT '12_2010_all.mat']  % USER Change --------------
matfile2 = [dir_MAT '01_2010_REGIONS_COR2.mat'] % USER Change ---------
%outfile1 = [dir_MAT '12_2010_all_for_EXCEL.txt'] % USER Change -------
outfile2 = [dir_MAT '01_2010_REGIONS_COR2_EXCEL.txt']  % USER Change --
% NOTE:  Number of  columns in FINAL_MATRIX MUST MATCH THE NUMBER OF
%14.4f\t  in the format statement
%  -------------- WRITE File ---------------------------
% load input MAT
feval('load', matfile2)

fid2 = fopen(outfile2,'w')
index = find(FINAL_MATRIX == Inf);
FINAL_MATRIX(index) = NaN;

for i=1:length(FINAL_MATRIX)
    one_row = FINAL_MATRIX(i,:)';
    %fprintf(fid2,'%14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t
%14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t
%14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t
%14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t
%14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t \n',one_row);
    % set for 30 columns
    fprintf(fid2,'%14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t
%14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t
%14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t
%14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t %14.4f\t
\n',one_row);

    clear one_row
end
fclose(fid2)
clear one_row FINAL_MATRIX
```

APPENDIX C. ADDITIONAL RESULTS FIGURES

This appendix provides additional figures for comparing of the various performance metrics between months and locations.

- Overall performance by month
- Hit proportion and level-1 and level-2 differences with respect to data latency issues by month
- Percent of monthly and yearly data by latitude
- Cloud category performance metric evaluations for all regions

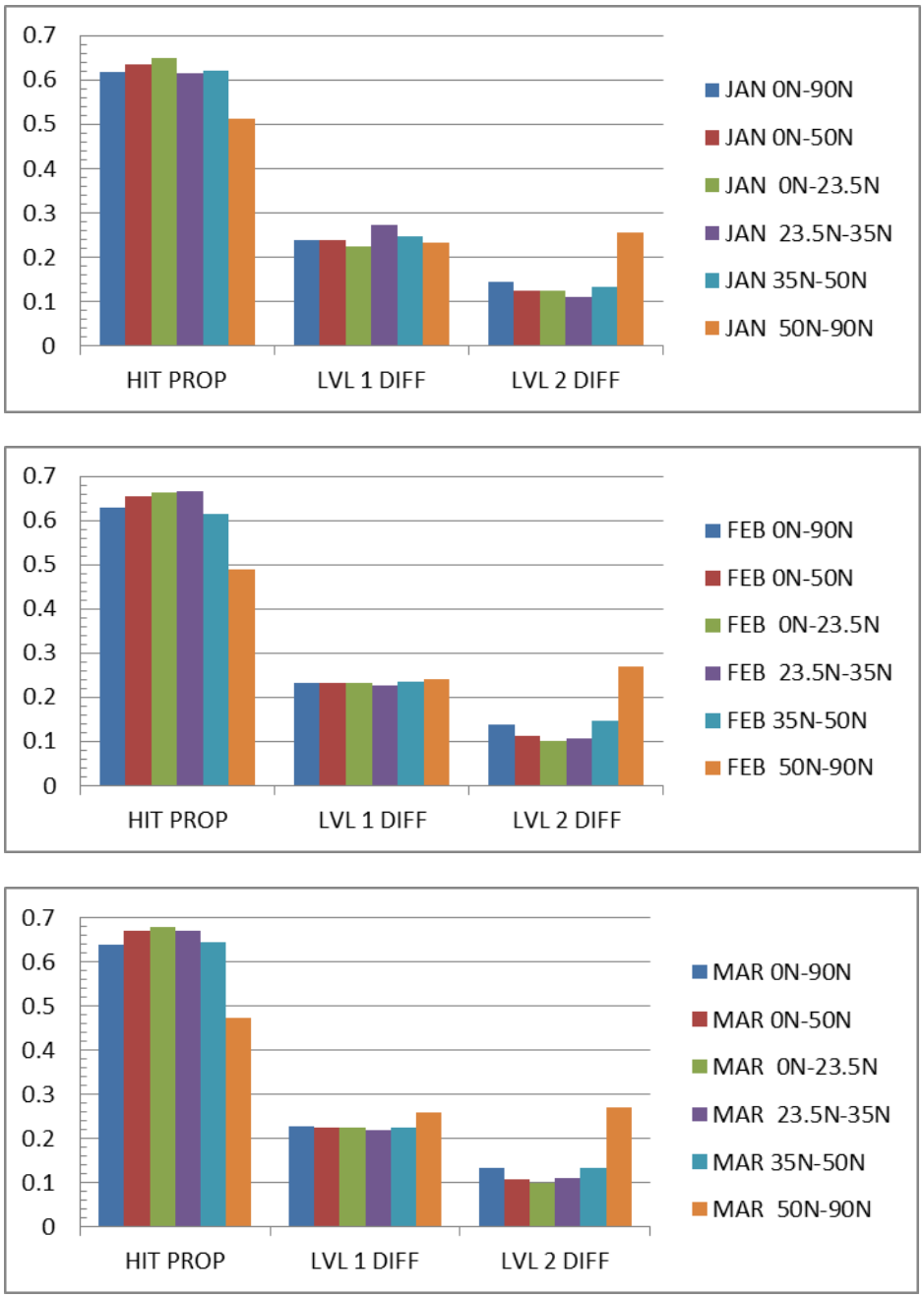


Figure 36. A comparative evaluation of WWMCA hit proportion and level-1 and level-2 differences for the six sub regions by month. From top to bottom the panels are January, February, and March. This figure should be compared to Figures 37–39.

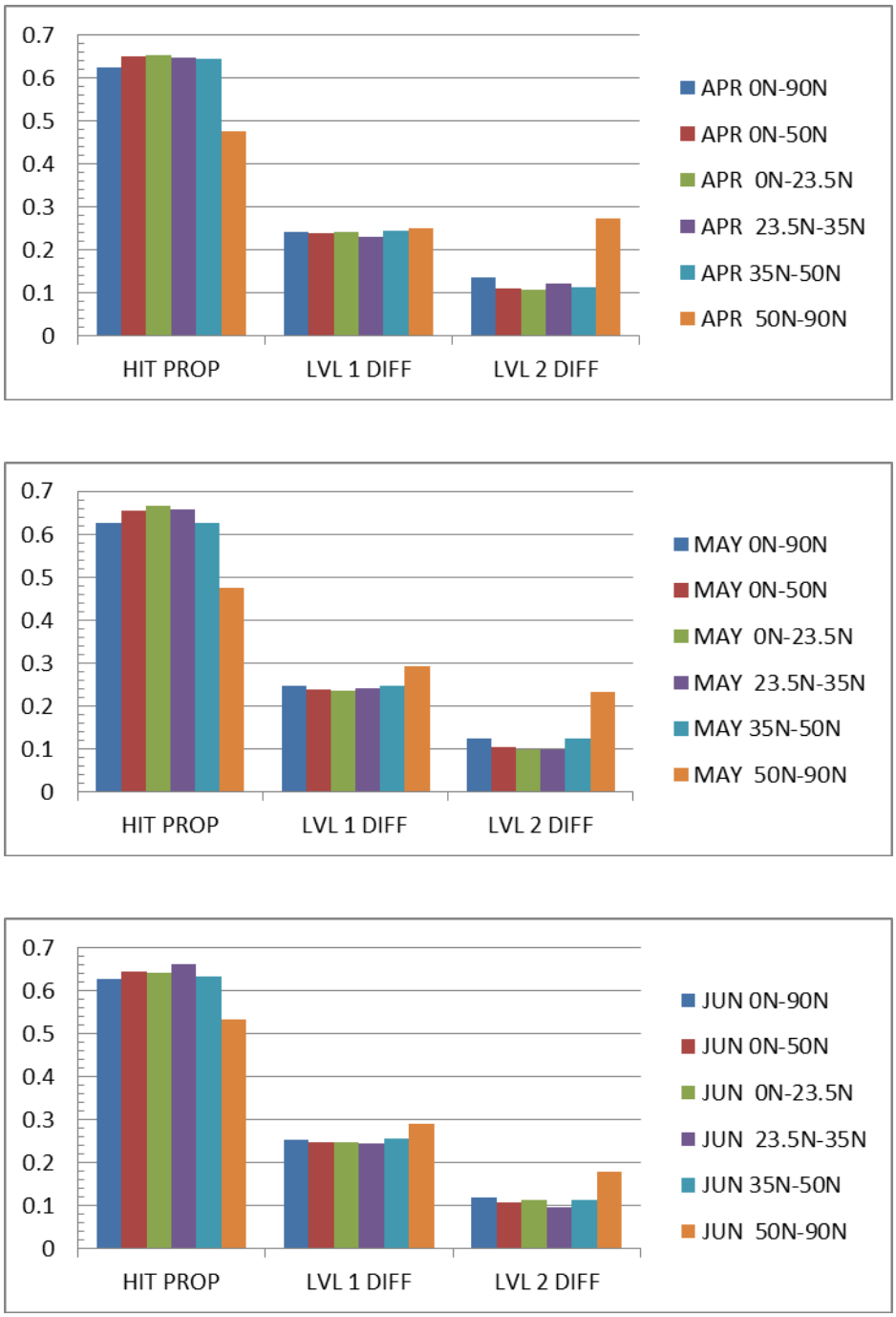


Figure 37. A comparative evaluation of WWMCA hit proportion and level-1 and level-2 differences for the six sub regions by month. From top to bottom the panels are April, May, and June. This figure should be compared to Figures 36, 38 and 39.

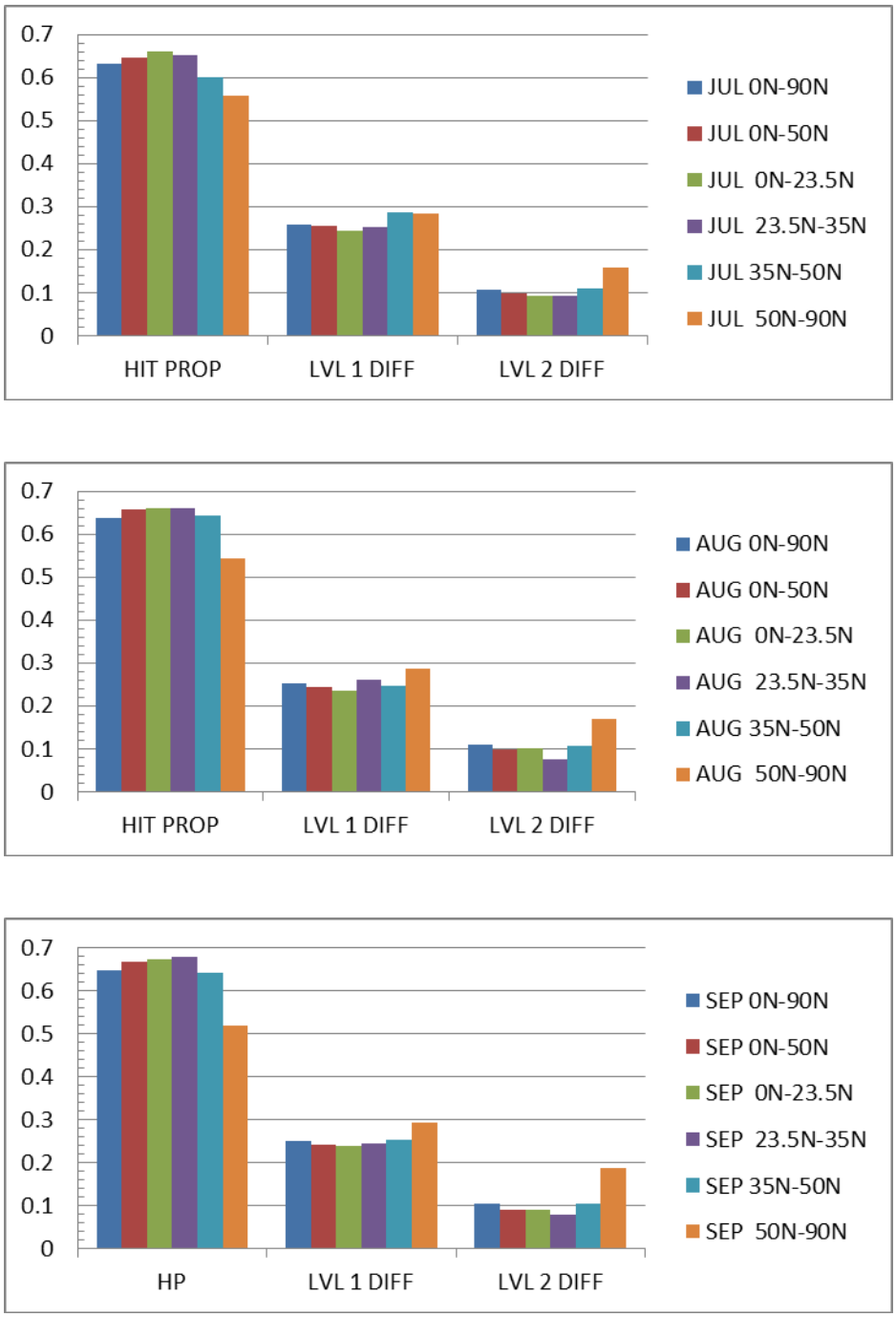


Figure 38. A comparative evaluation of WWMCA hit proportion and level-1 and level-2 differences for the six sub regions by month. From top to bottom the panels are July, August, and September. This figure should be compared to Figures 35, 36 and 39.

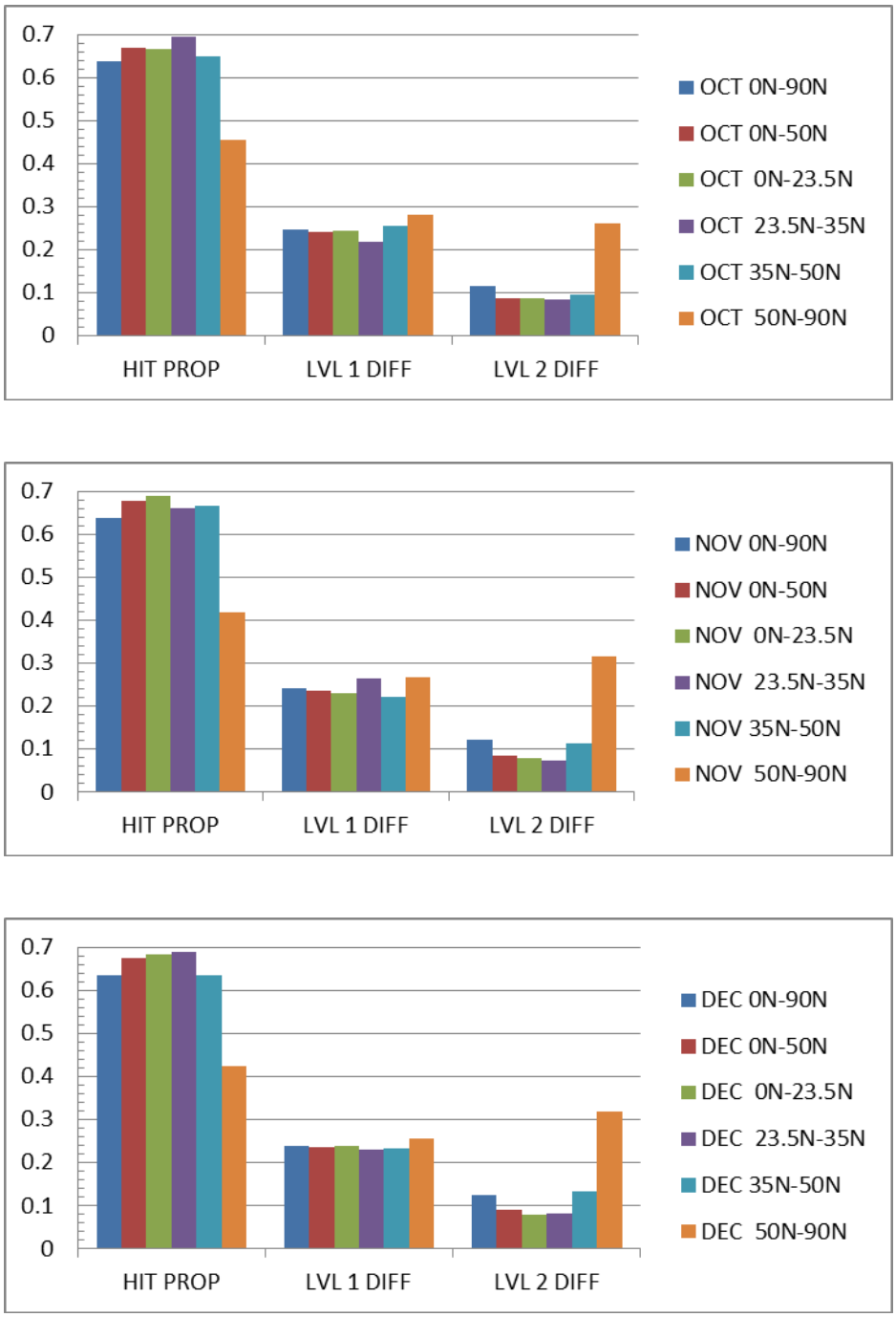


Figure 39. A comparative evaluation of WWMCA hit proportion and level-1 and level-2 differences for the six sub regions by month. From top to bottom the panels are October, November, and December. This figure should be compared to Figures 36–38.

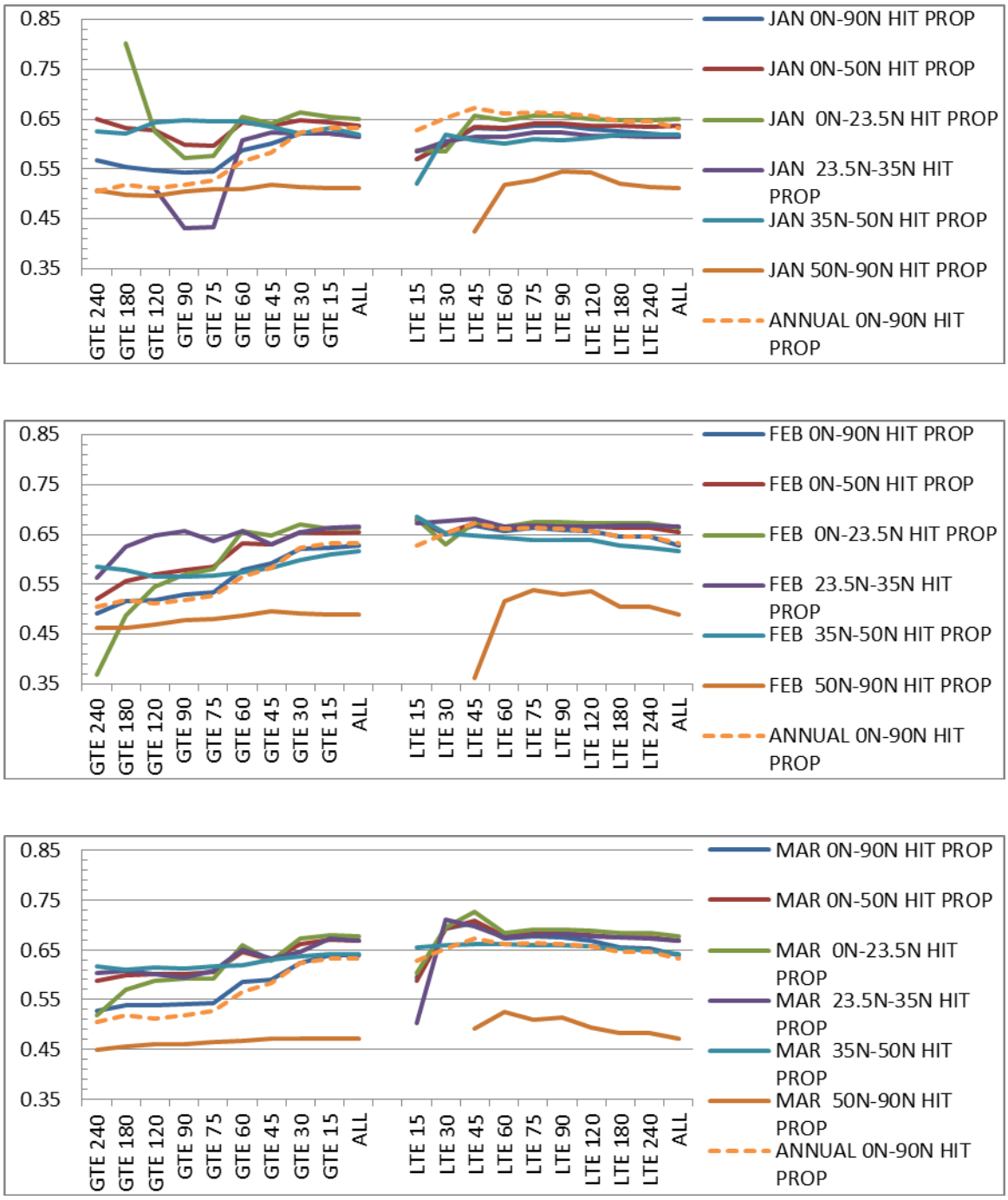


Figure 40. Monthly hit proportion for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are January, February, and March. This figure should be compared to Figures 41 – 43.

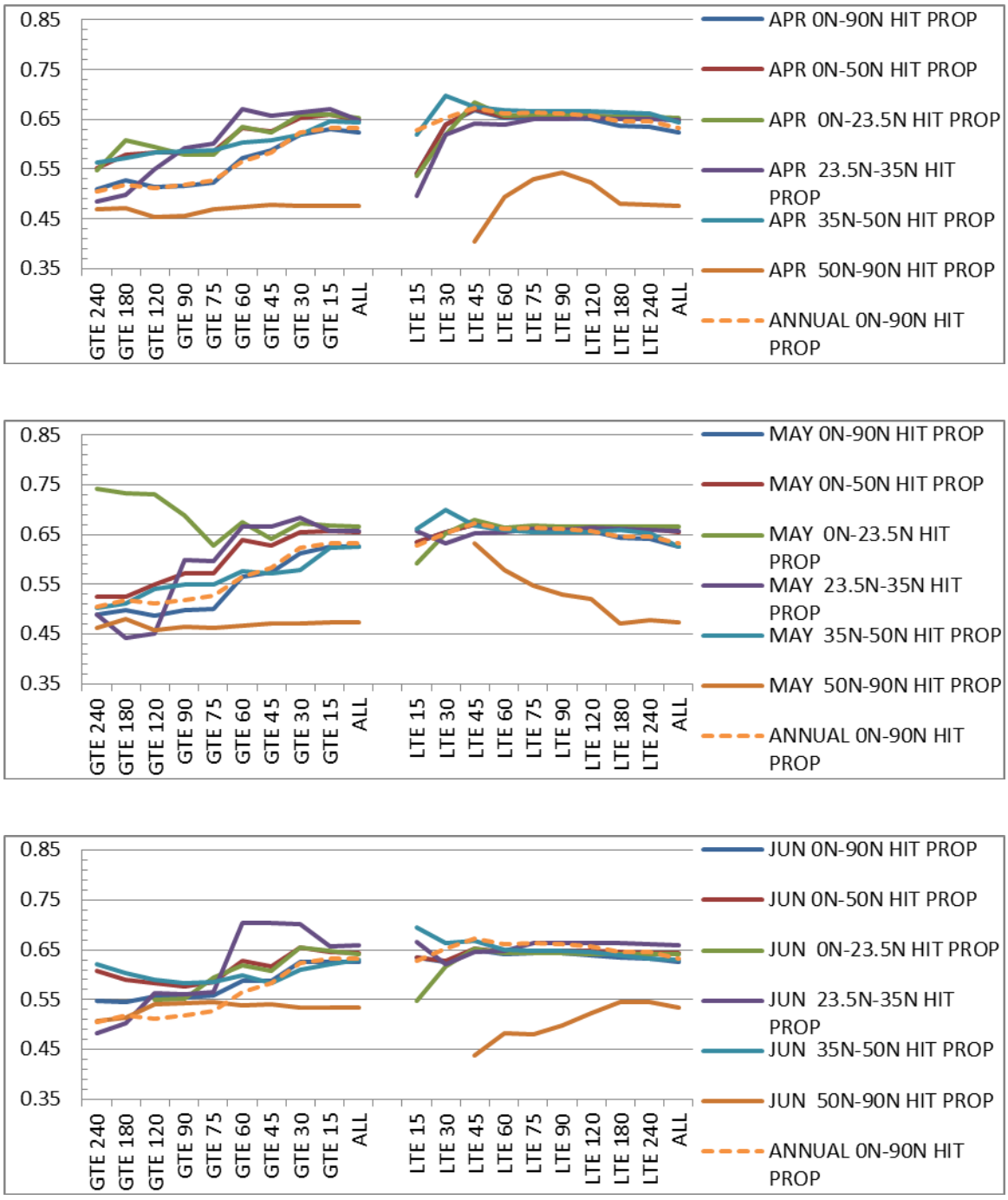


Figure 41. Monthly hit proportion correct for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are April, May, and June. This figure should be compared to Figures 40, 42 and 43.

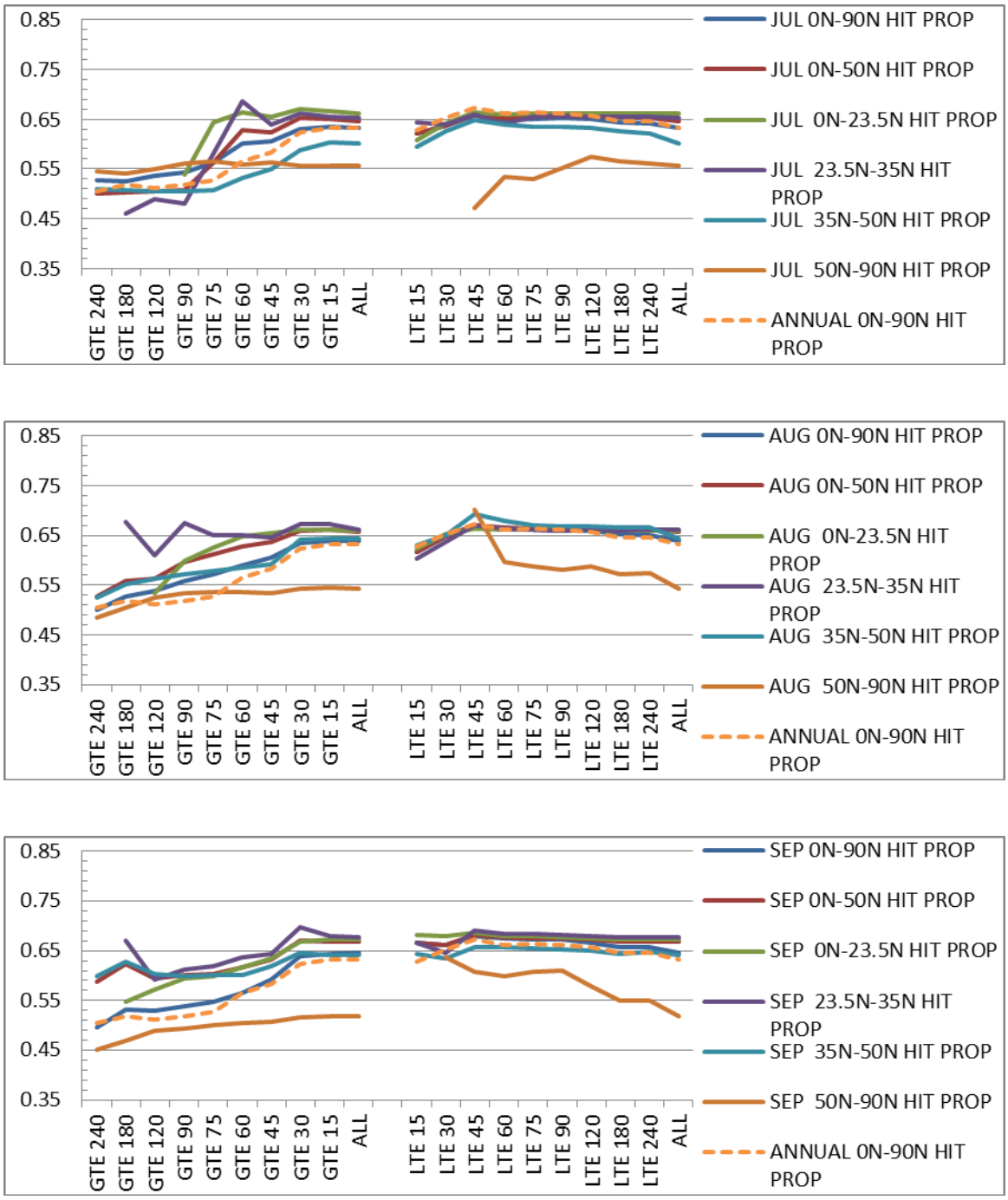


Figure 42. Monthly hit proportion for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are July, August, and September. This figure should be compared to Figures 40, 41 and 43.

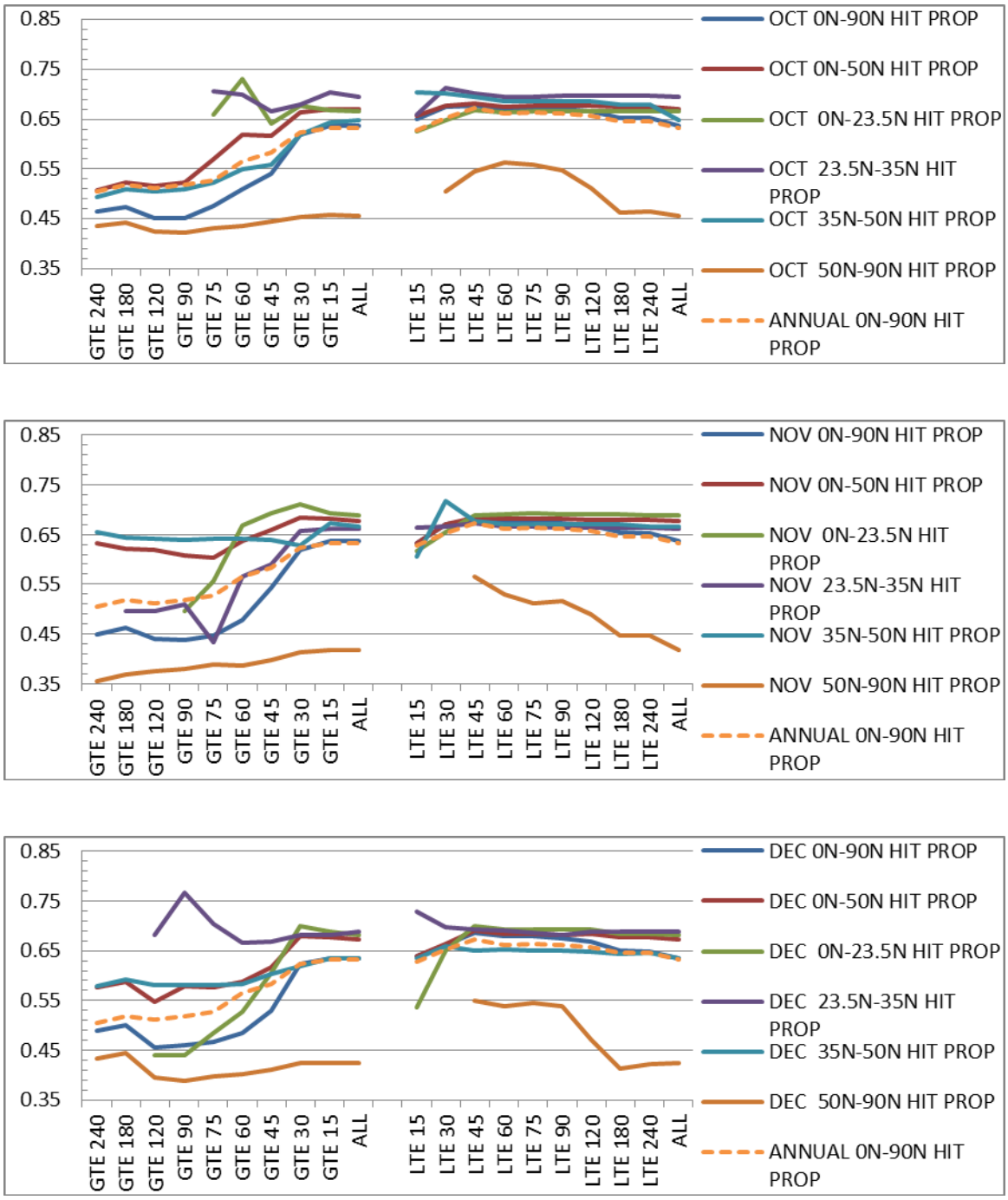


Figure 43. Monthly hit proportion for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are October, November, and December. This figure should be compared to Figures 40 – 42.

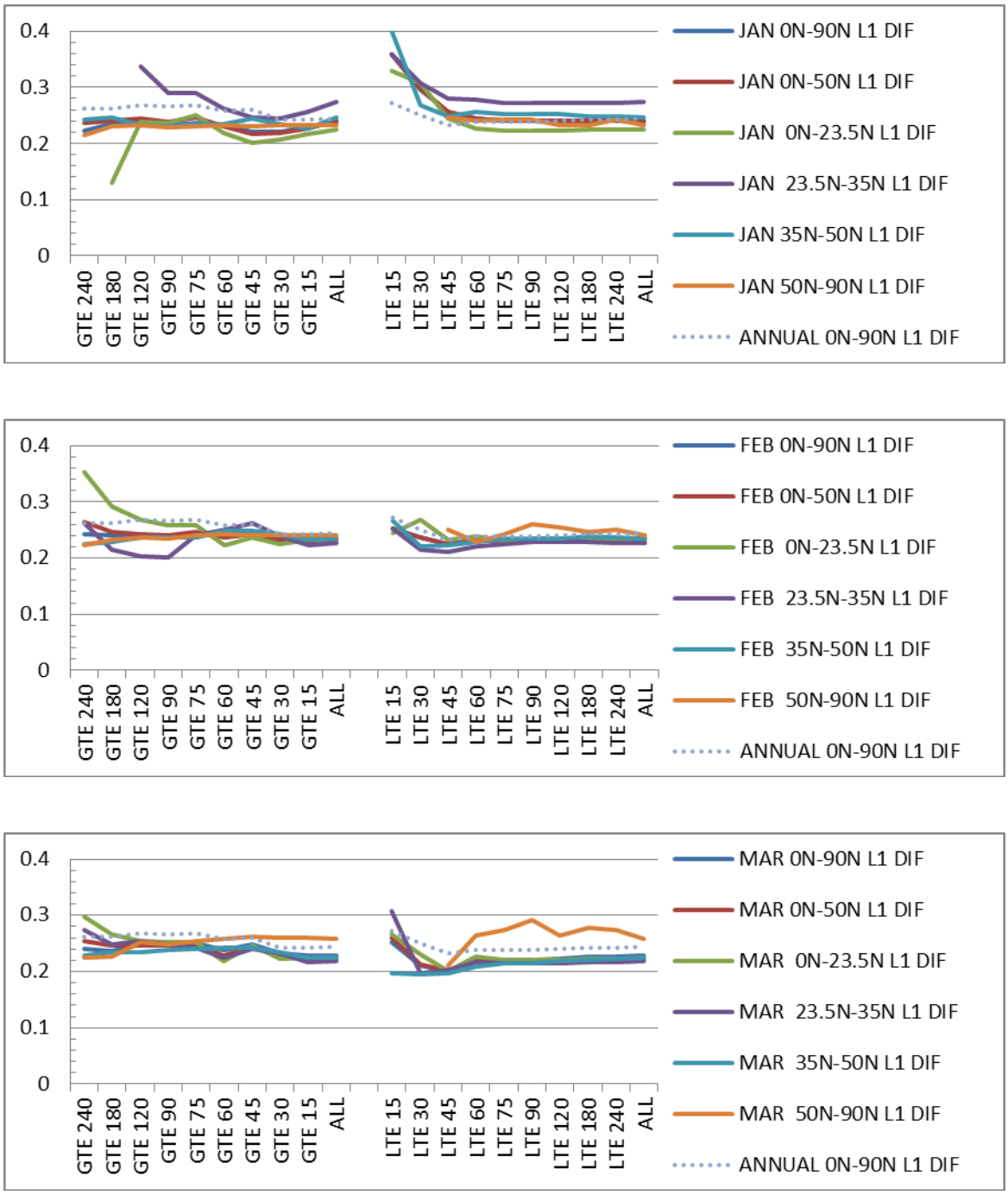


Figure 44. Monthly level-1 differences for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are January, February, and March. This figure should be compared to Figures 45 – 47.

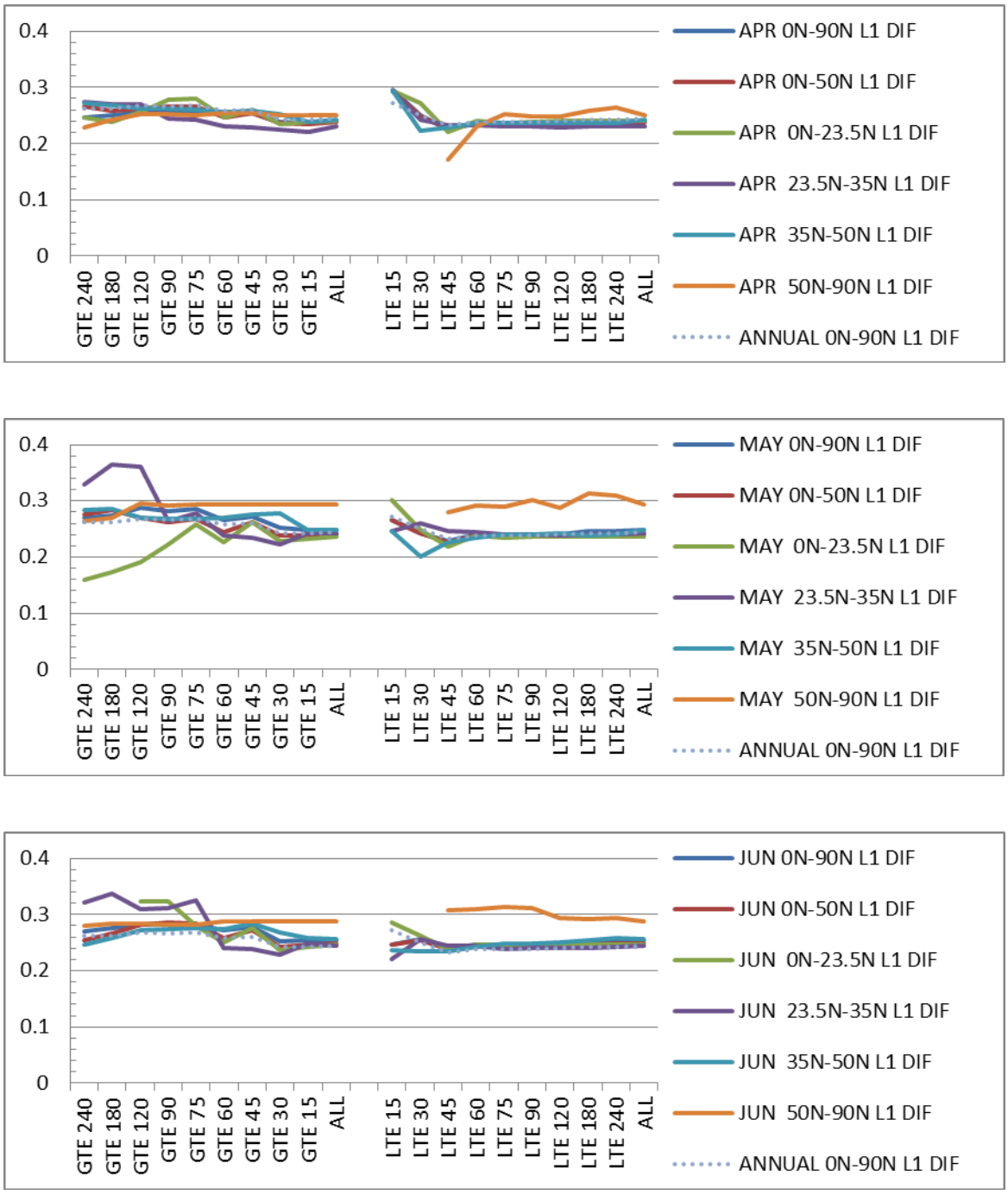


Figure 45. Monthly level-1 differences for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are April, May, and June. This figure should be compared to Figures 44, 46, 47.

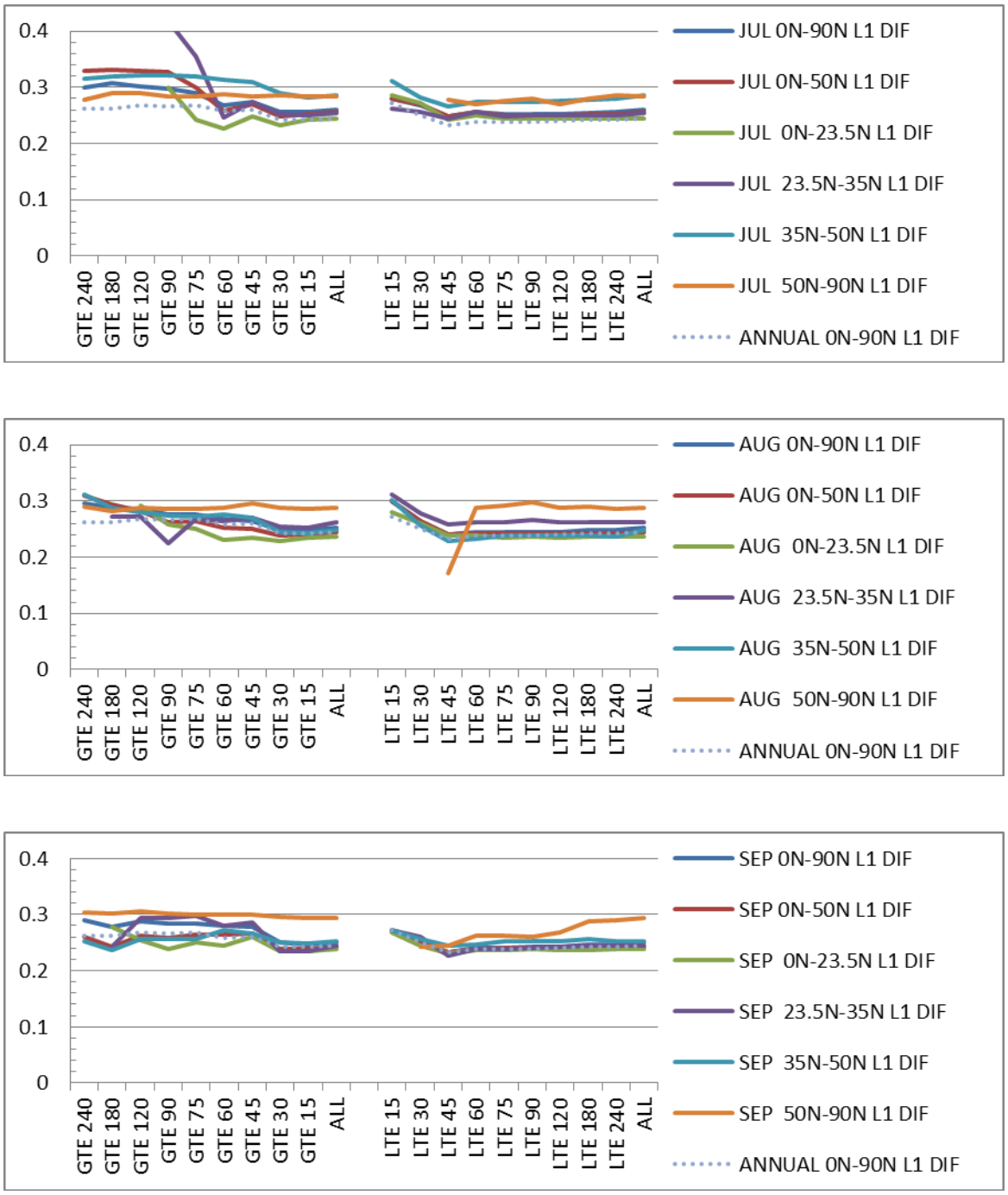


Figure 46. Monthly level-1 differences for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are July, August, and September. This figure should be compared to Figures 44, 45 and 47.

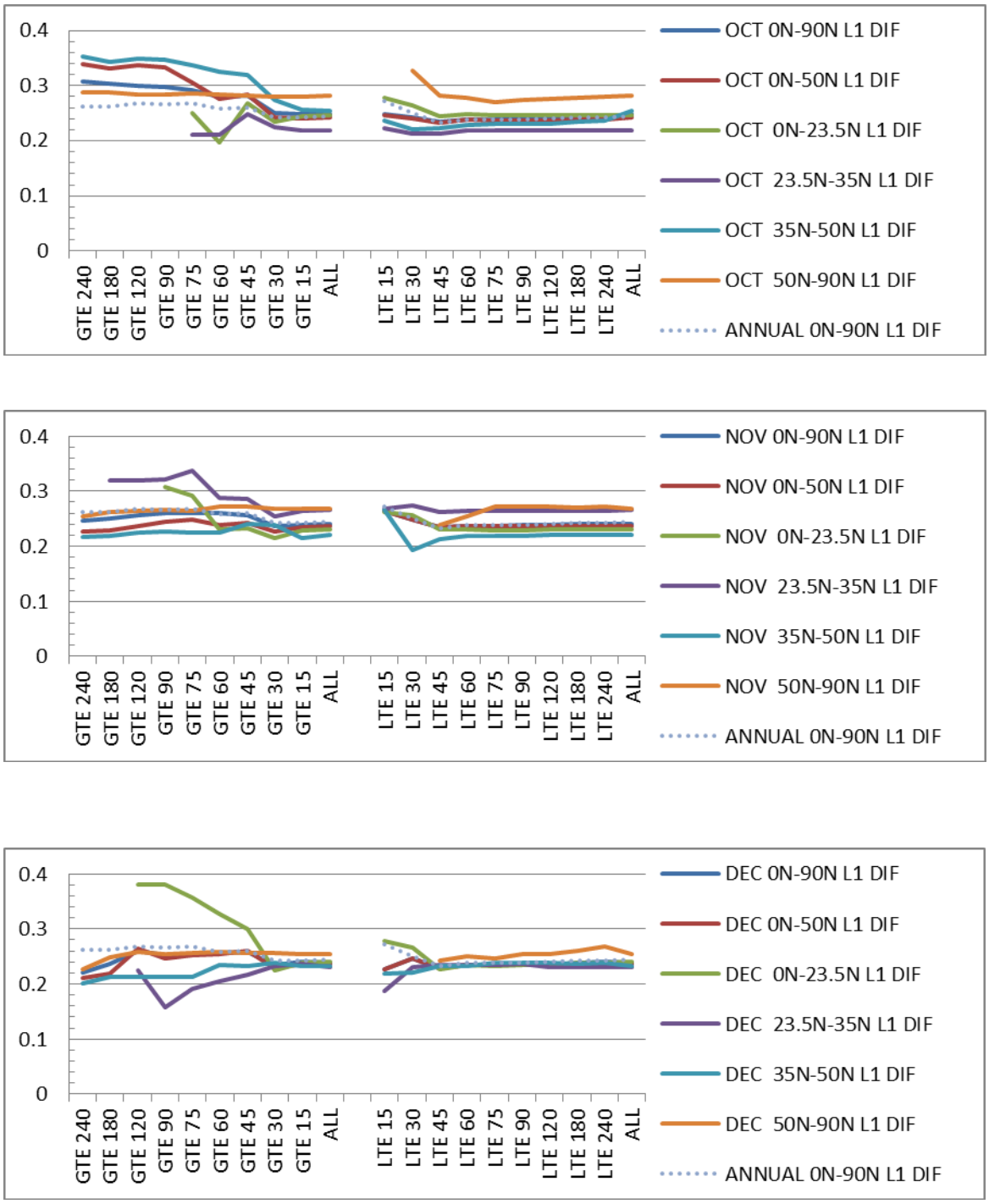
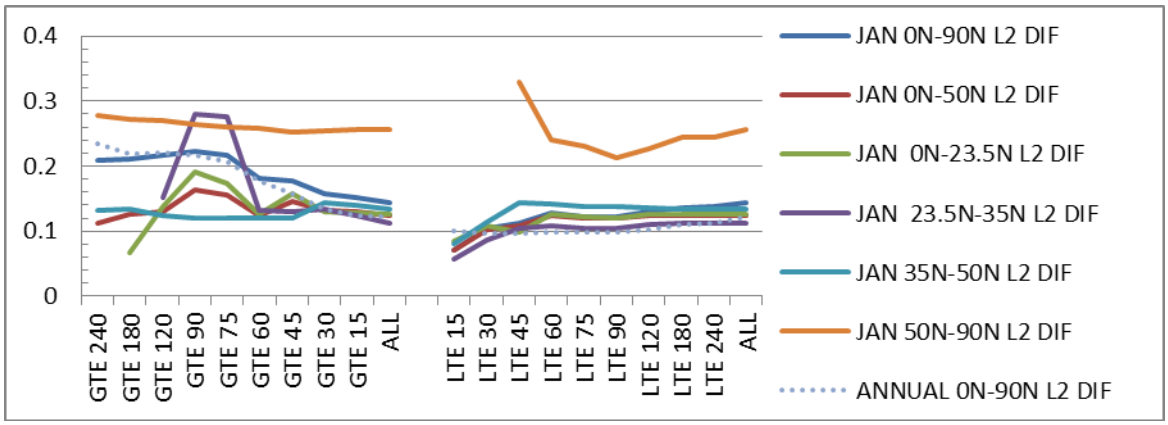
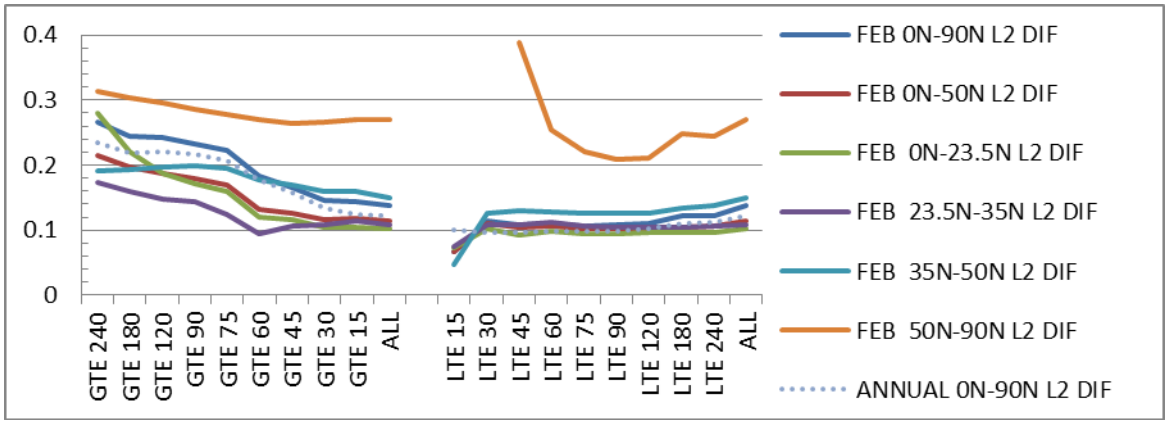


Figure 47. Monthly level-1 differences for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are October, November, and December. This figure should be compared to Figures 44 – 46.





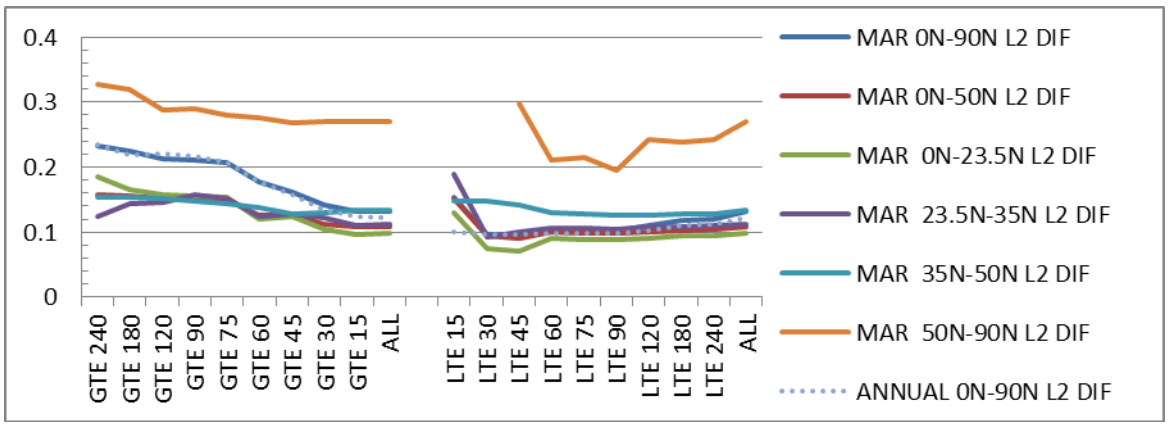


Figure 48. Monthly level-2 differences for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are January, February, and March. This figure should be compared to Figures 49 – 51.

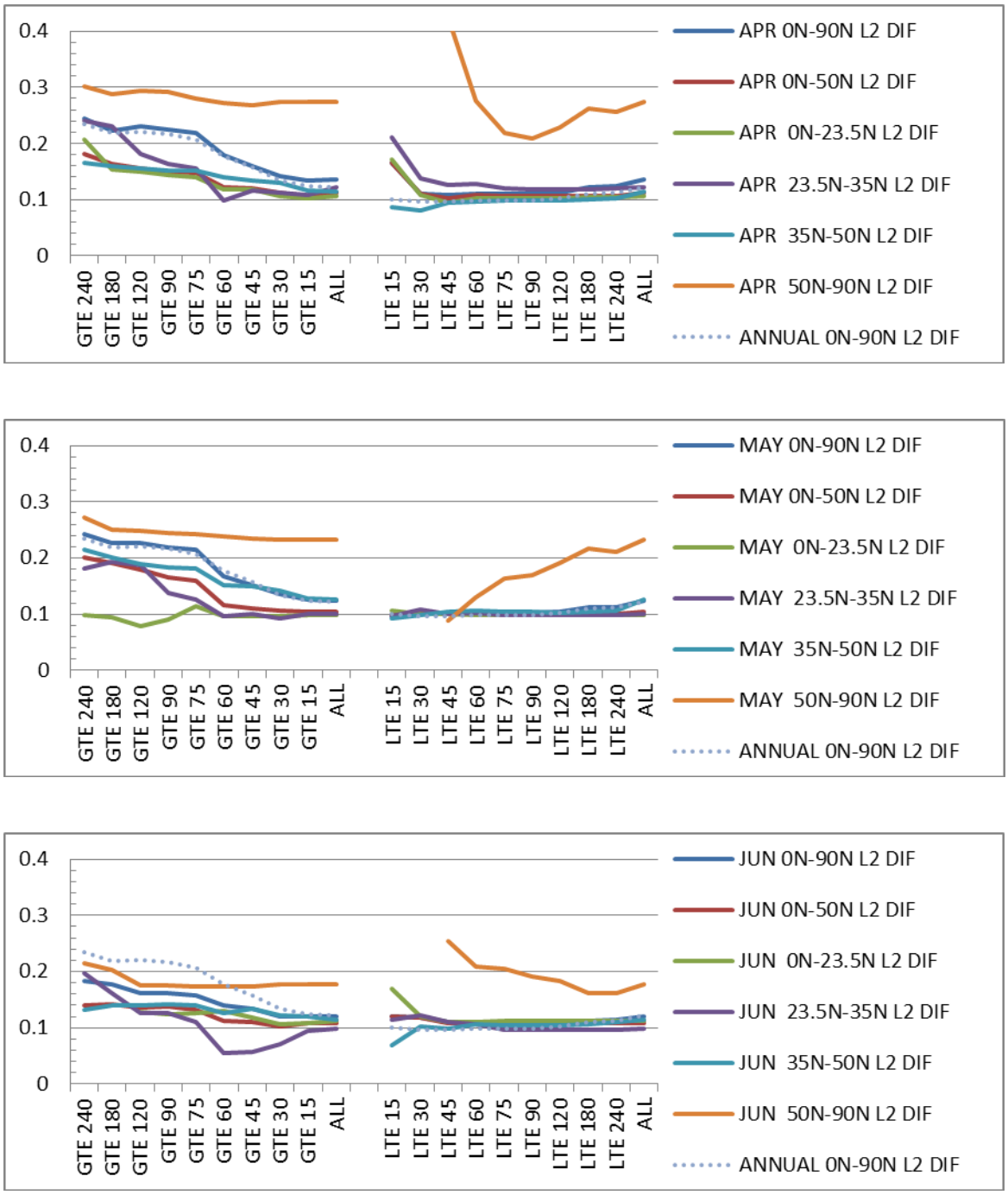


Figure 49. Monthly level-2 differences for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are April, May, and June. This figure should be compared to Figures 48, 50, and 51.

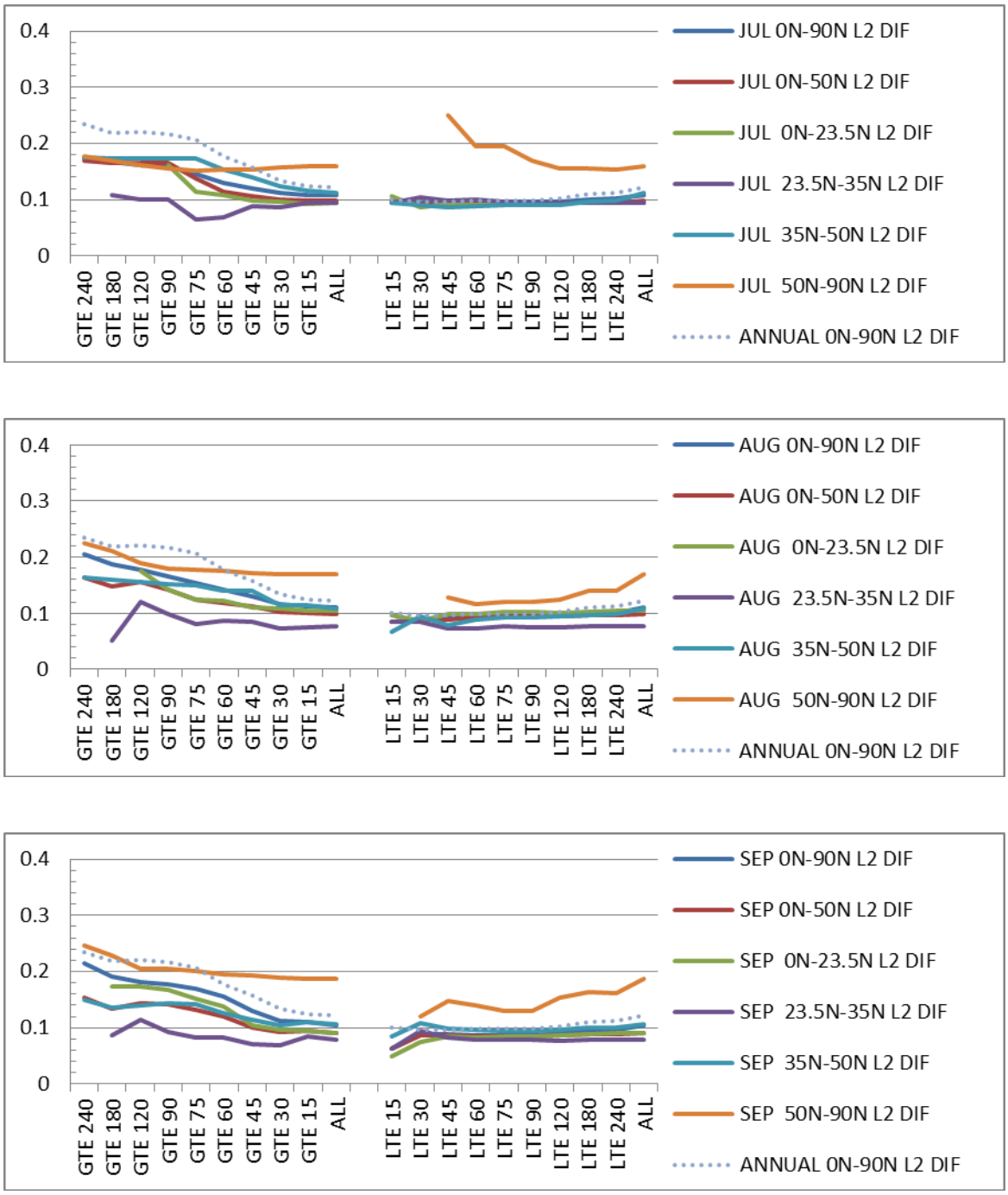


Figure 50. Monthly level-2 differences for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are July, August, and September. This figure should be compared to Figures 48, 49 and 51.

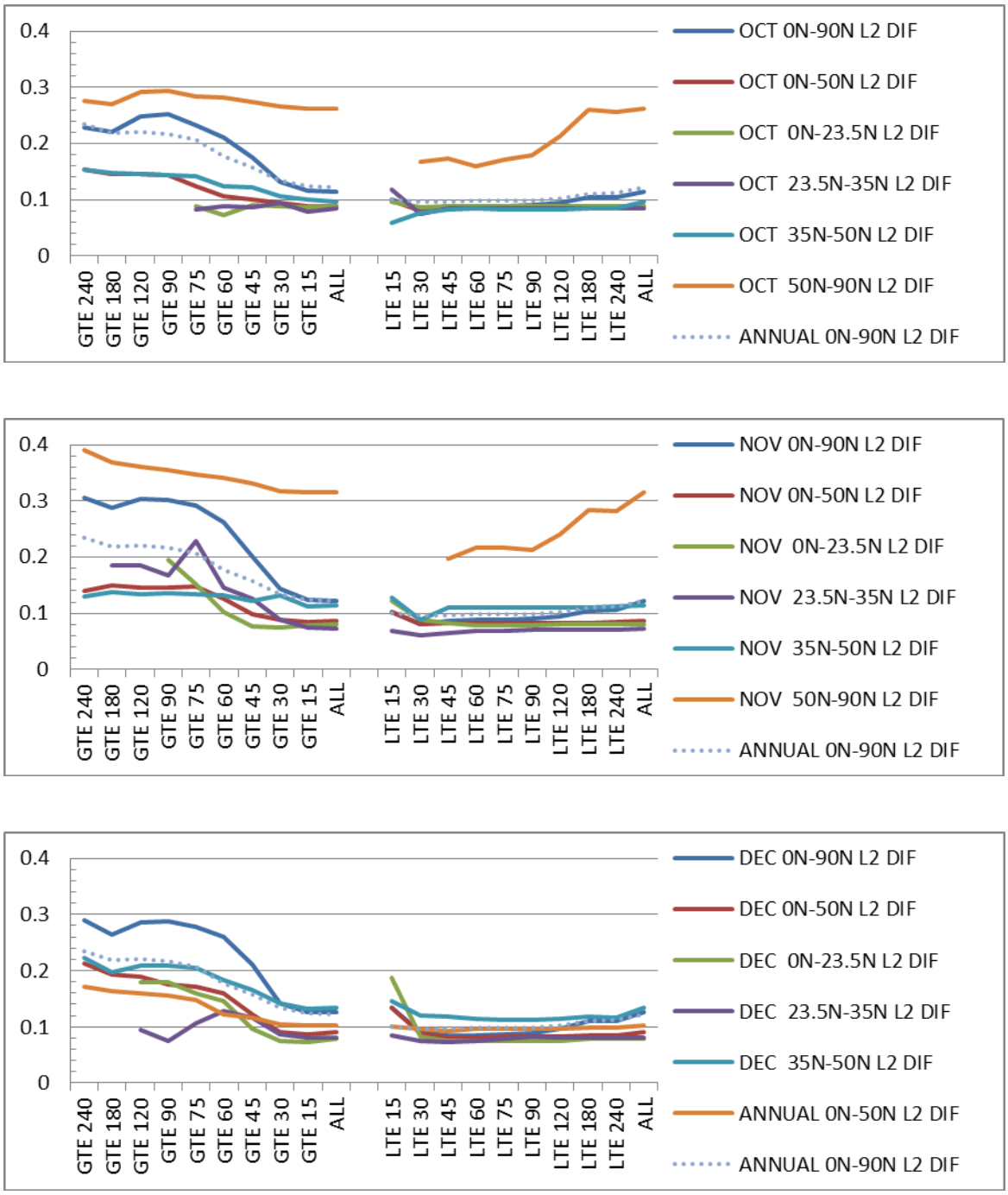


Figure 51. Monthly level-2 differences for each region in comparison to the annual WWMCA performance for all time categories. From top to bottom these panels are October, November, and December. This figure should be compared to Figures 48 – 50.

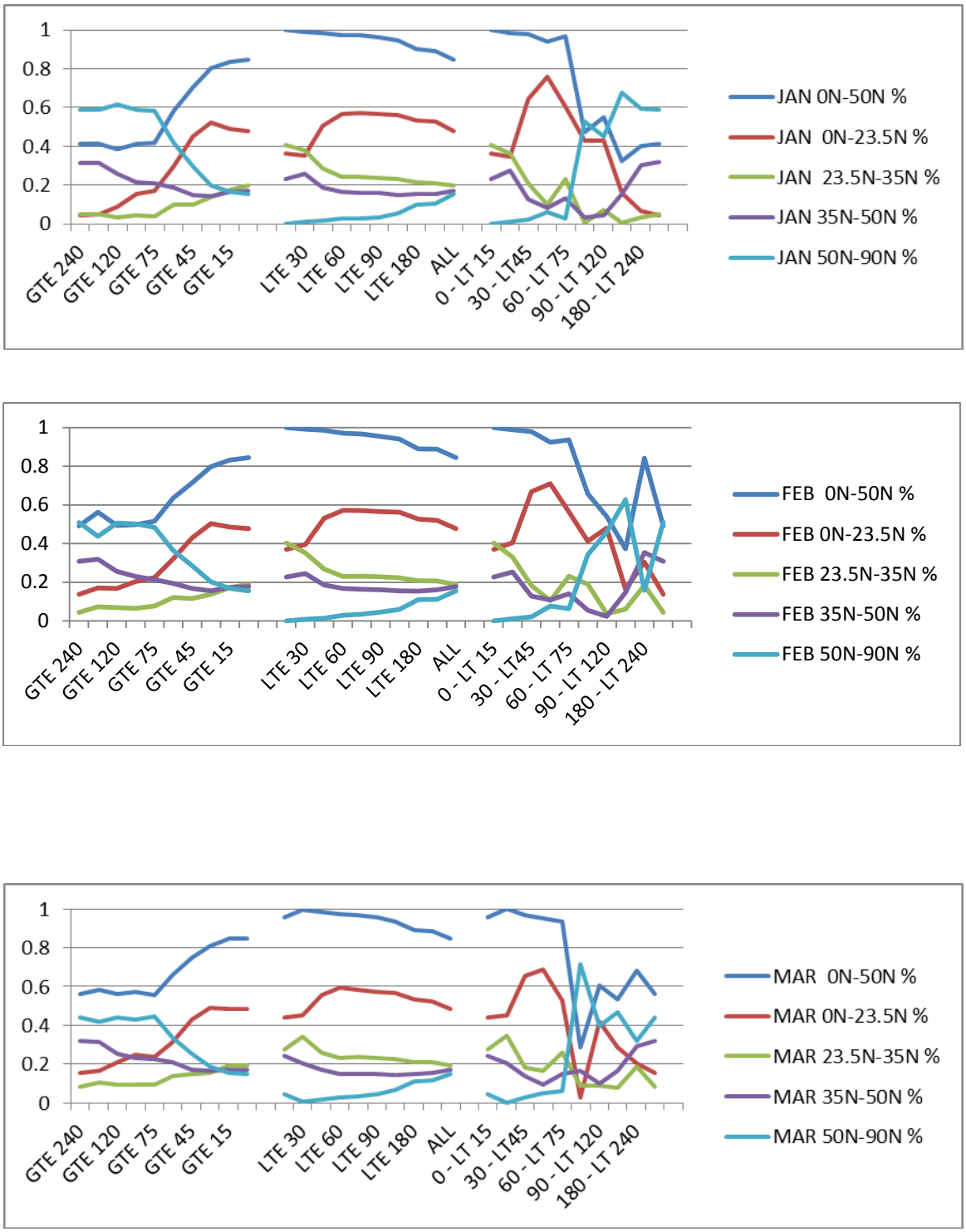
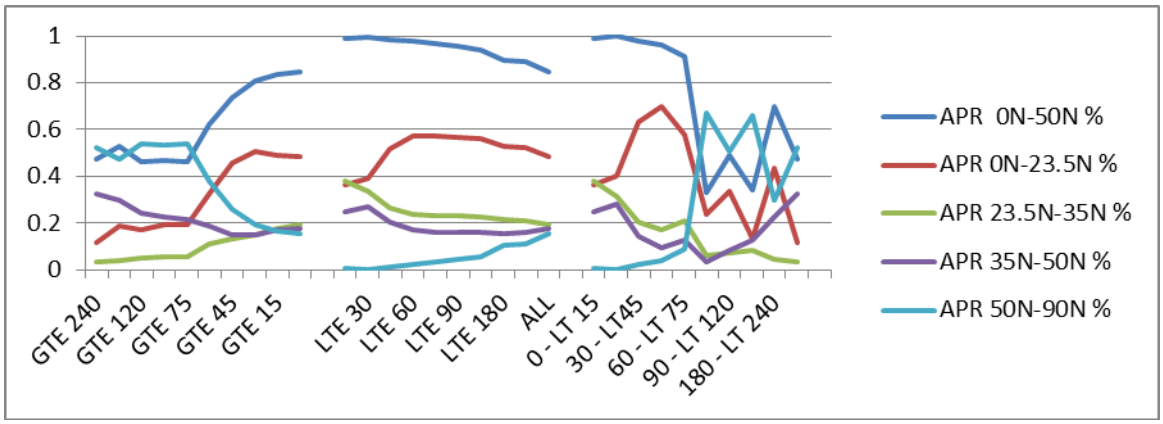
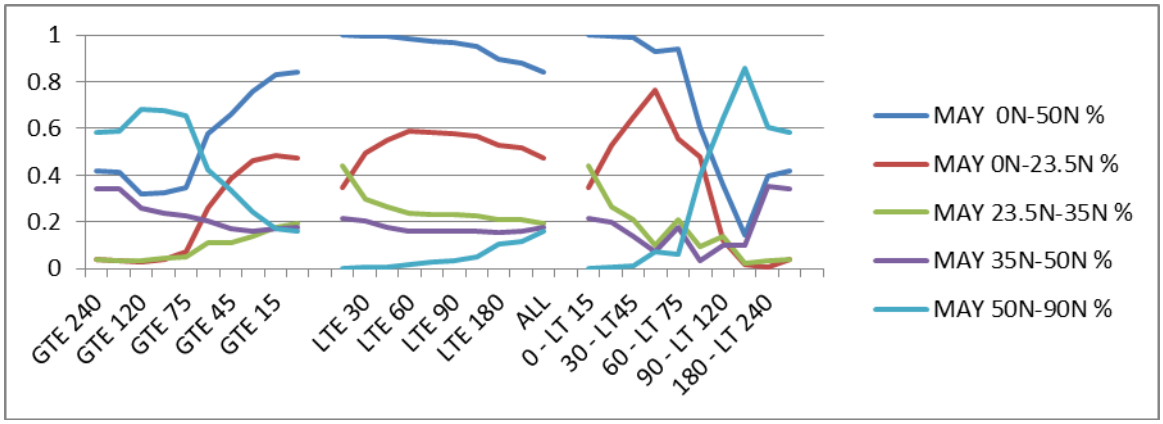


Figure 52. Monthly percent of total data coverage by latitude. Value represents the percent of data out of all data collected for that period. From top to bottom these panels are January, February, and March. This figure should be compared to Figures 53 – 55.





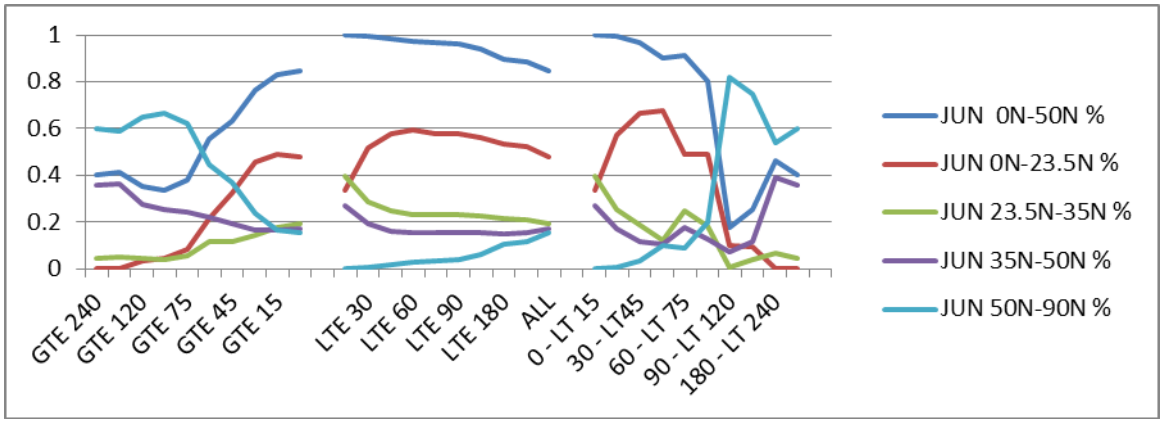


Figure 53. Monthly percent of total data coverage by latitude. Value represents the percent of data out of all data collected for that period. From top to bottom these panels are April, May, and June. This figure should be compared to Figures 52, 54, and 55.

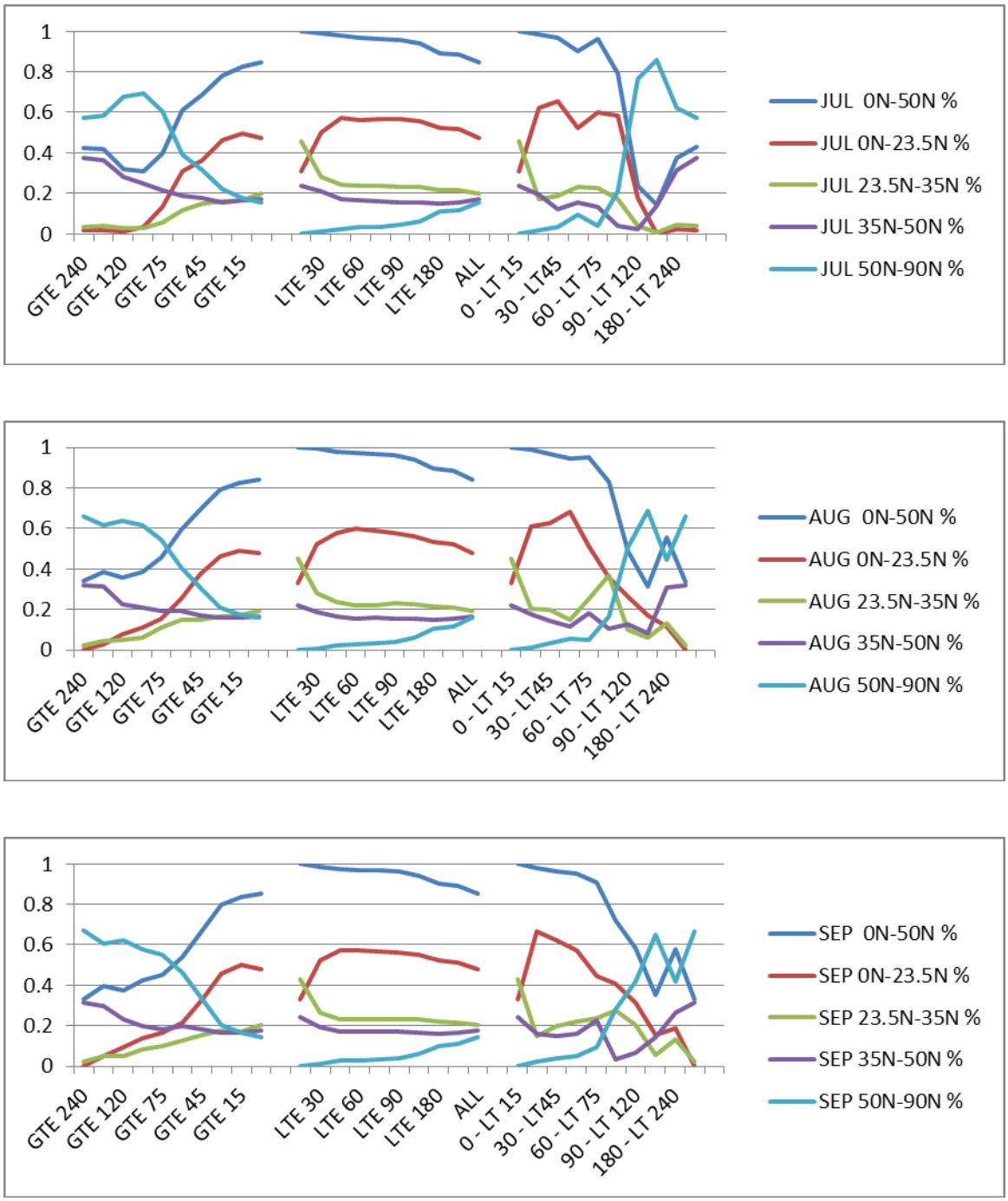


Figure 54. Monthly percent of total data coverage by latitude. Value represents the percent of data out of all data collected for that period. From top to bottom these panels are July, August, and September. This figure should be compared to Figures 52, 53, and 55.

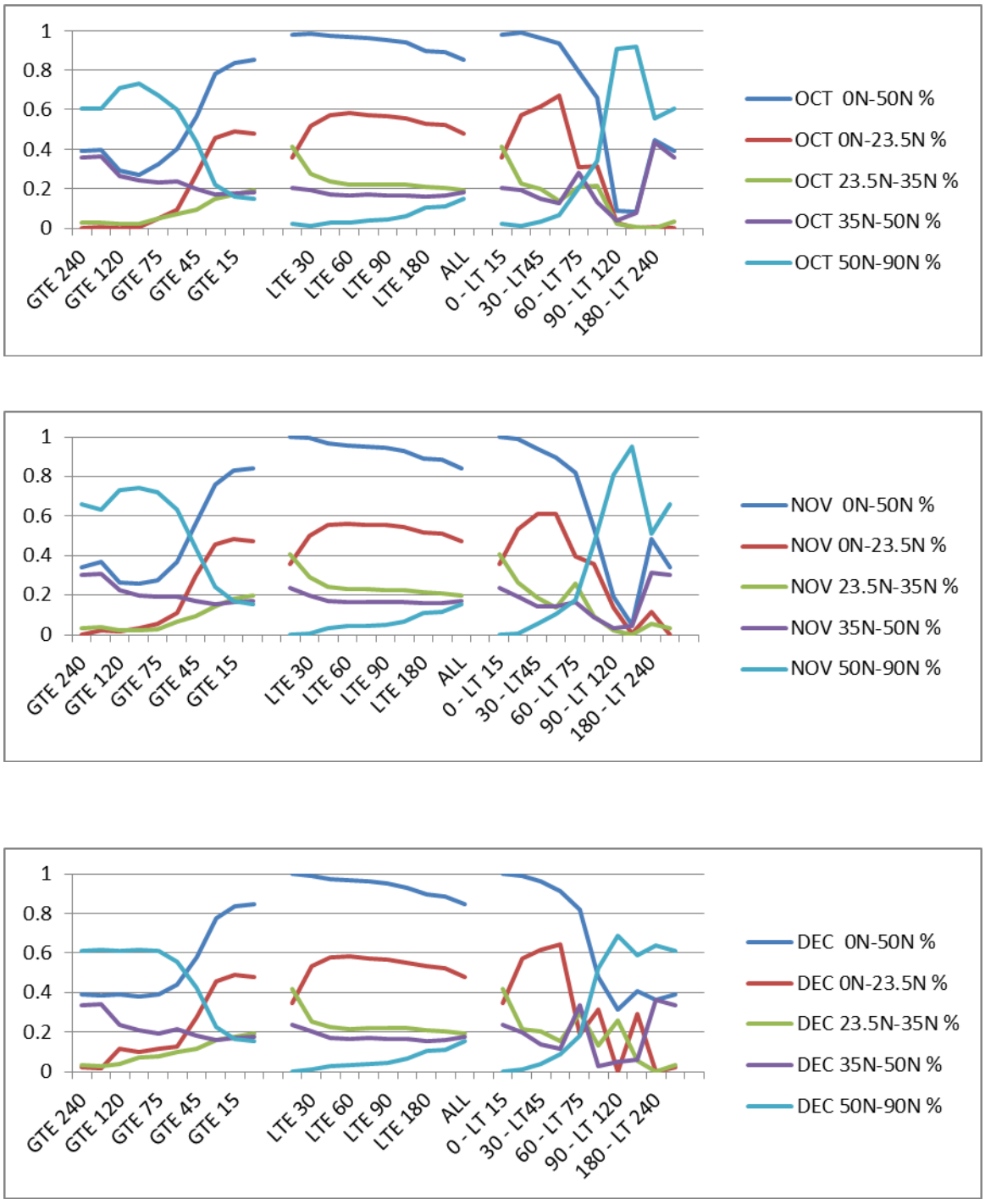


Figure 55. Monthly percent of total data coverage by latitude. Value represents the percent of data out of all data collected for that period. From top to bottom these panels are October, November, and December. This figure should be compared to Figures 52 – 54.

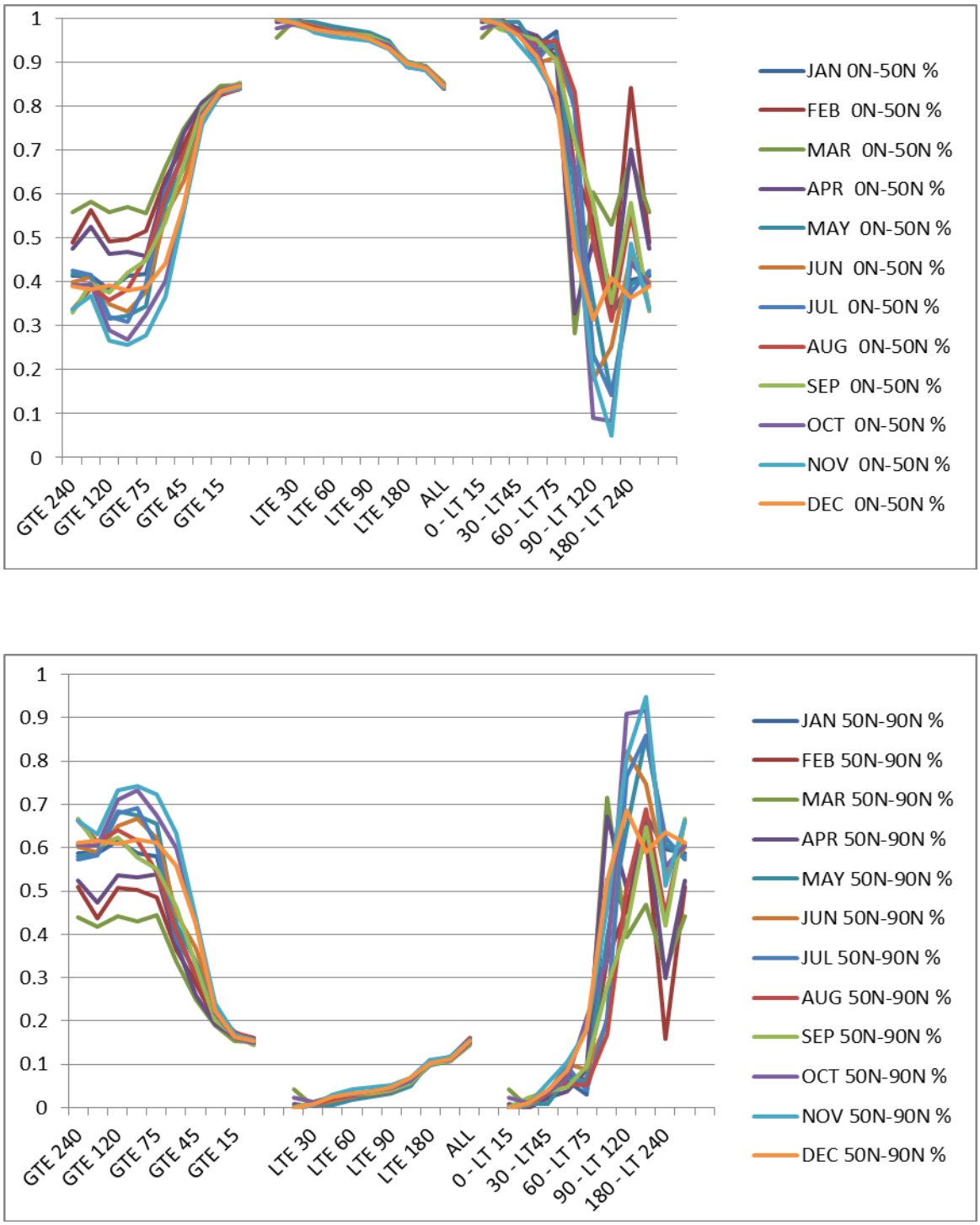


Figure 56. Annual variation of total data coverage by latitude. Value represents the percent of data out of all data collected for that period. The top panel is 0N–50N the bottom panel is 50N–90N. From left to right, the panels represent greater than or equal to minutes, less than or equal to minutes and minute ranges. This figure should be compared to Figures 57 and 58.

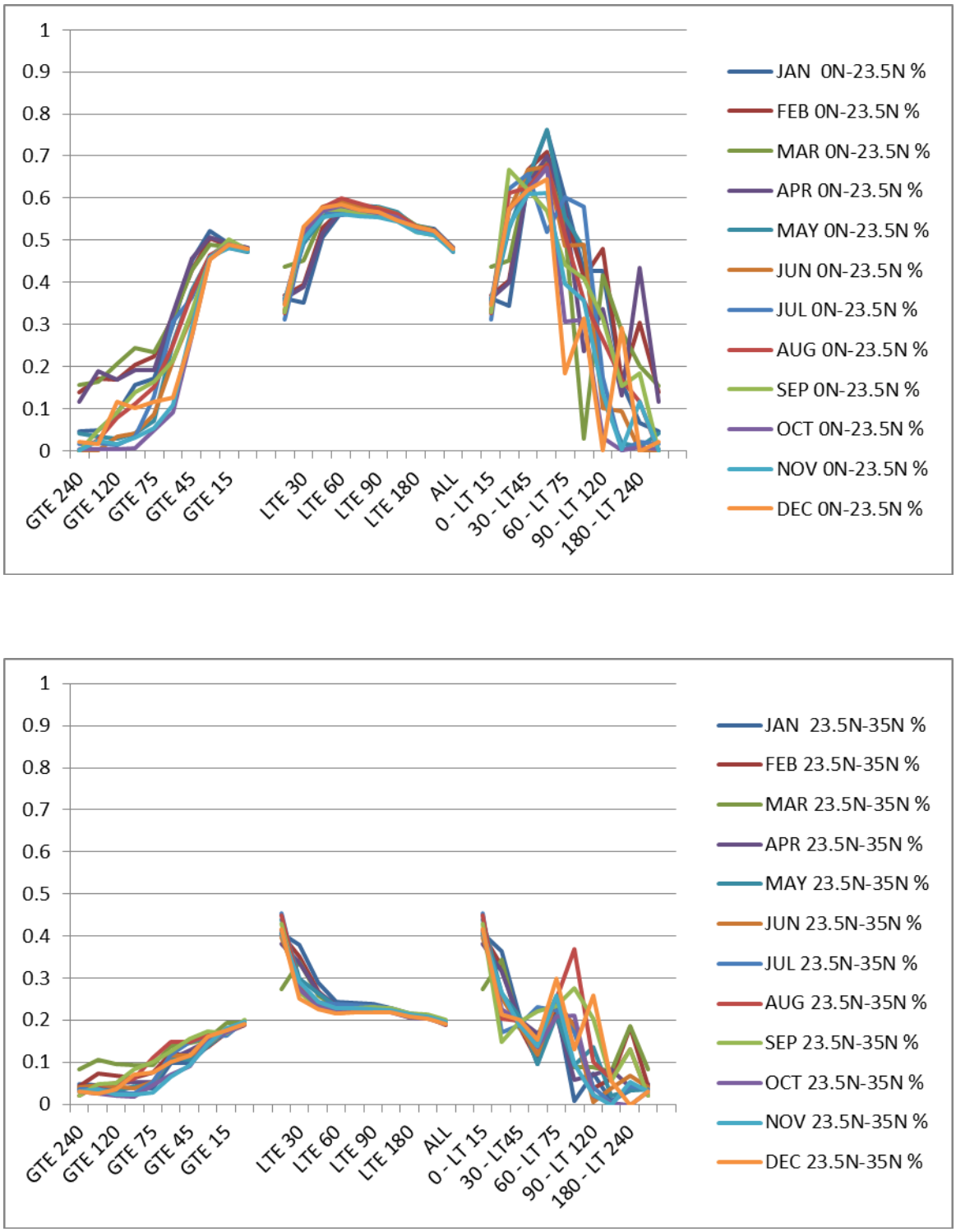


Figure 57. Annual variation of total data coverage by latitude. Value represents the percent of data out of all data collected for that period. The top panel is 0N–23.5N the bottom panel is 23.5N–35N. From left to right, the panels represent greater than or equal to minutes, less than or equal to minutes and minute ranges. This figure should be compared to Figures 56 and 58.

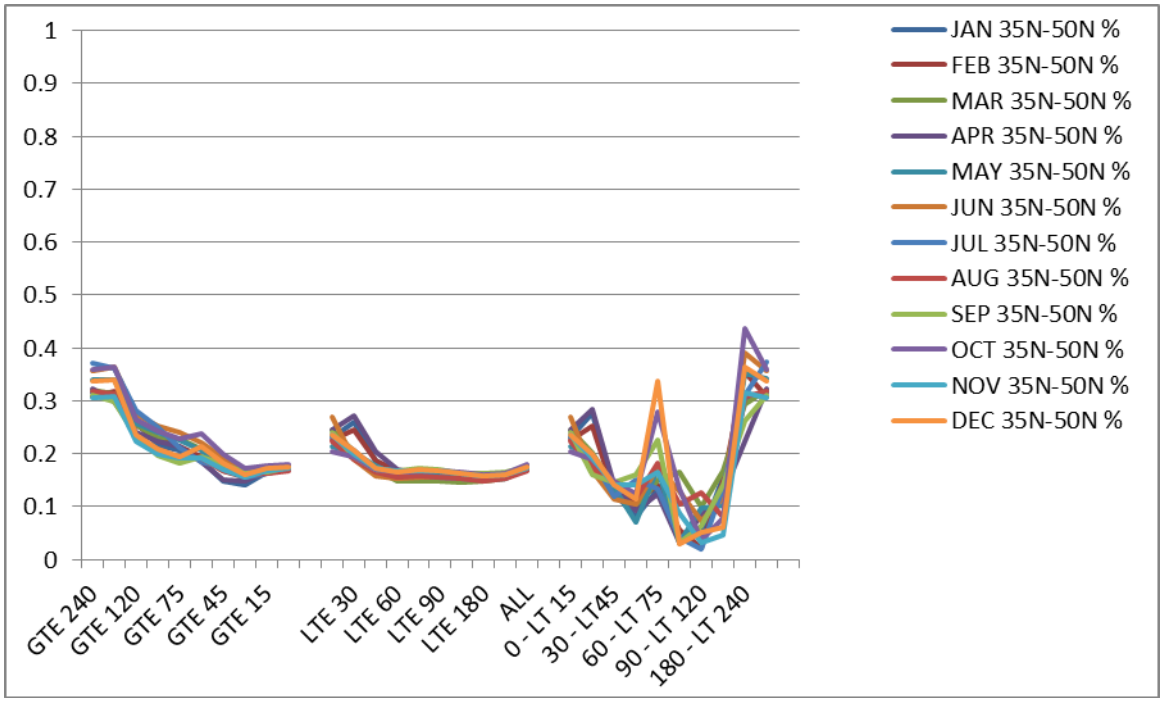


Figure 58. Annual variation of total data coverage by latitude. Value represents the percent of data out of all data collected for that period. This is 35N–50N. From left to right, the panels represent greater than or equal to minutes, less than or equal to minutes and minute ranges. This figure should be compared to Figures 56 and 57.

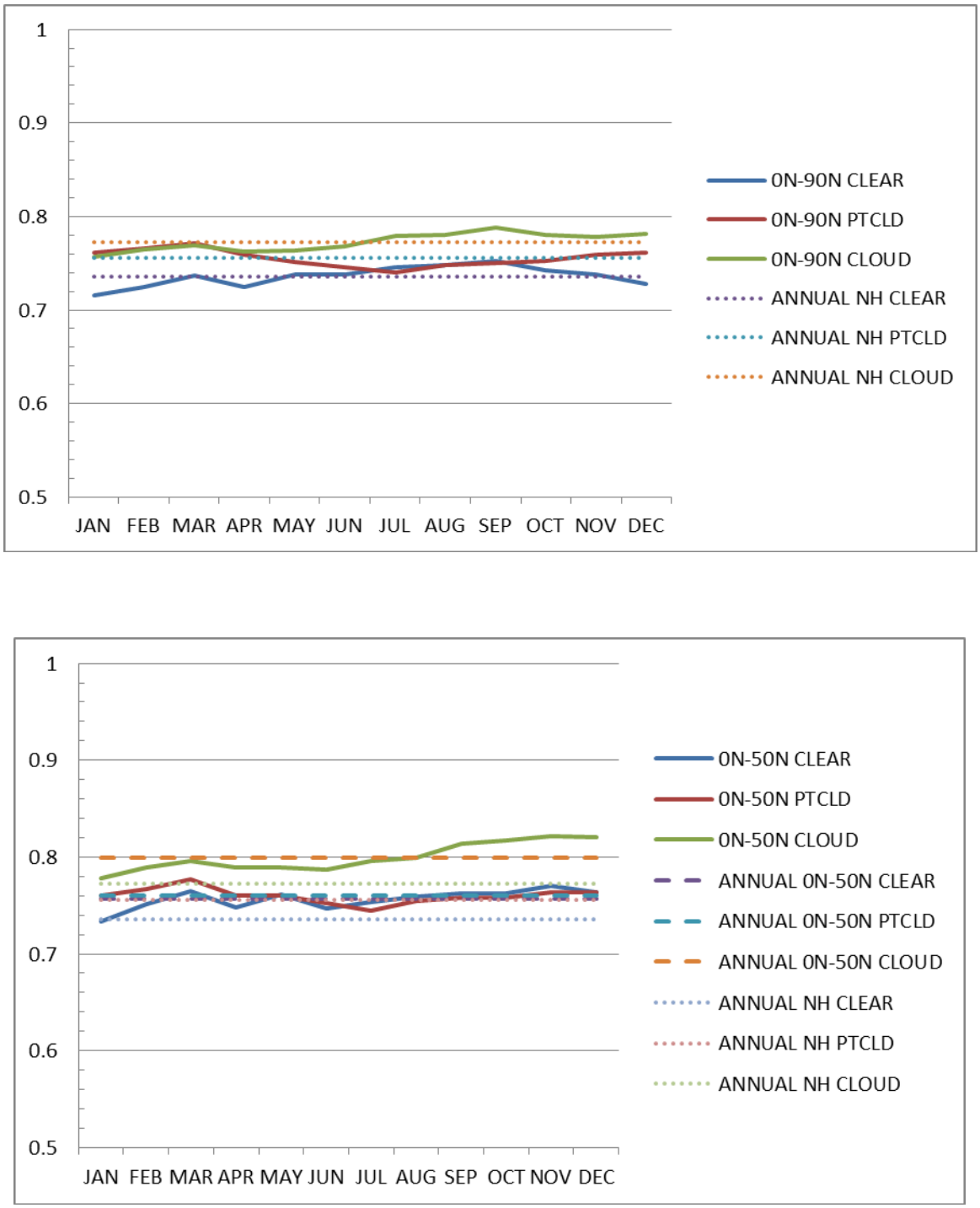


Figure 59. Proportion Correct for each month and all three cloud categories. The top panel is for 0N–90N while the bottom panel is for 0N–50N. This figure should be compared to Figures 60–62. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

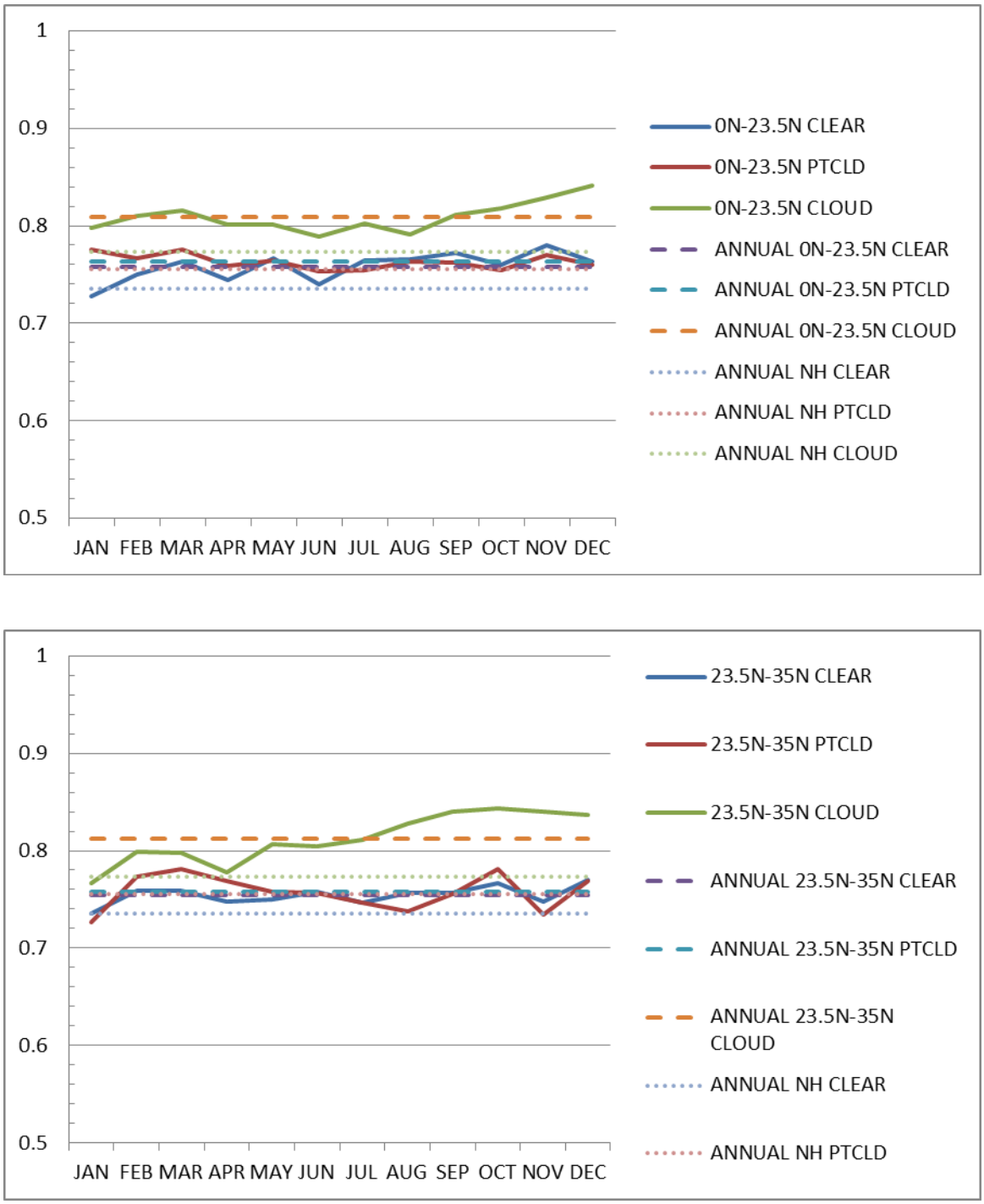


Figure 60. Proportion Correct for each month and all three cloud categories. The top panel is for 0N–23.5N while the bottom panel is for 23.5N–35N. This figure should be compared to Figures 59, 61, and 62. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

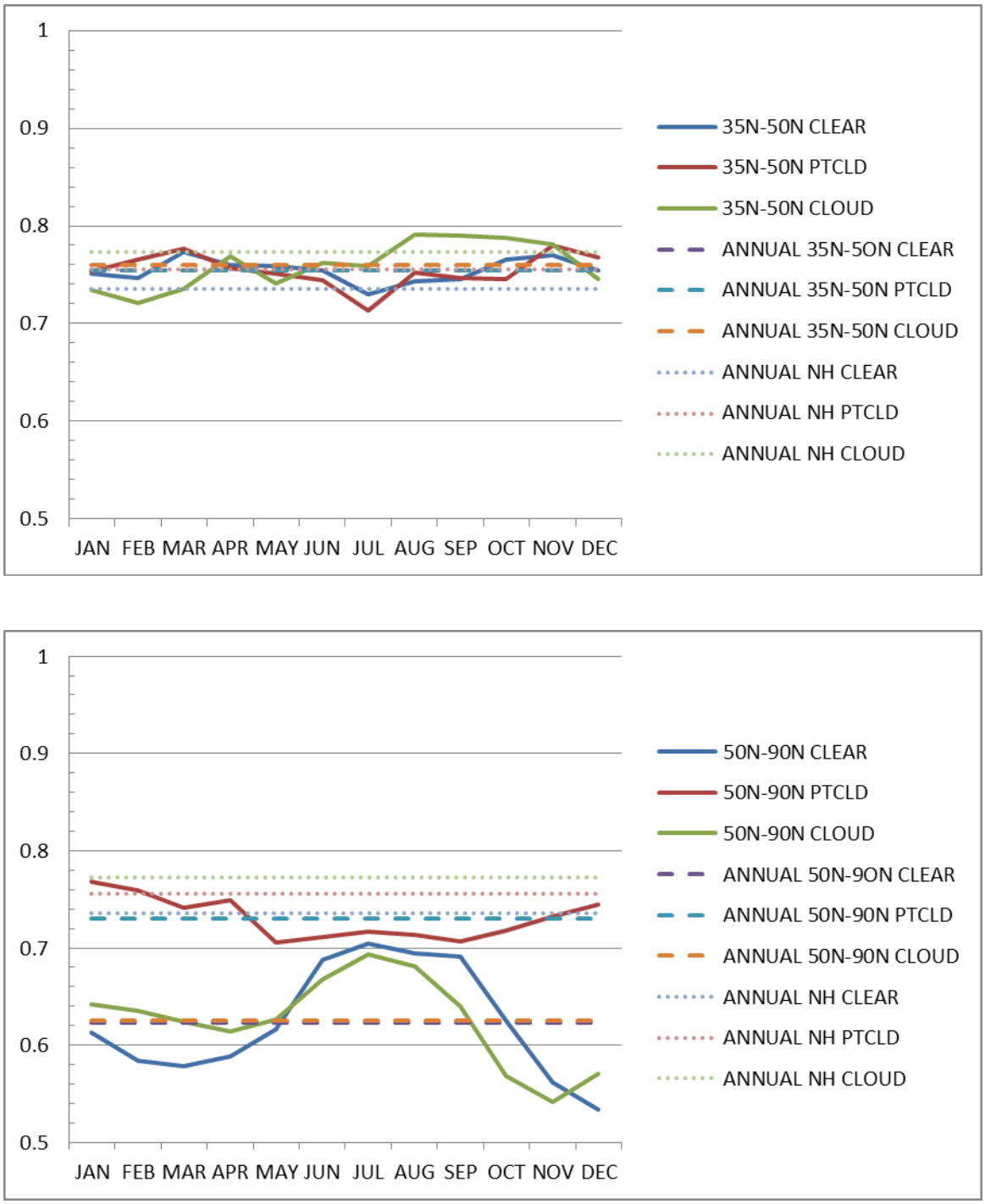


Figure 61. Proportion Correct for each month and all three cloud categories. The top panel is for 35N–50N while the bottom panel is for 50N–90N. This figure should be compared to Figures 59, 60, and 62. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

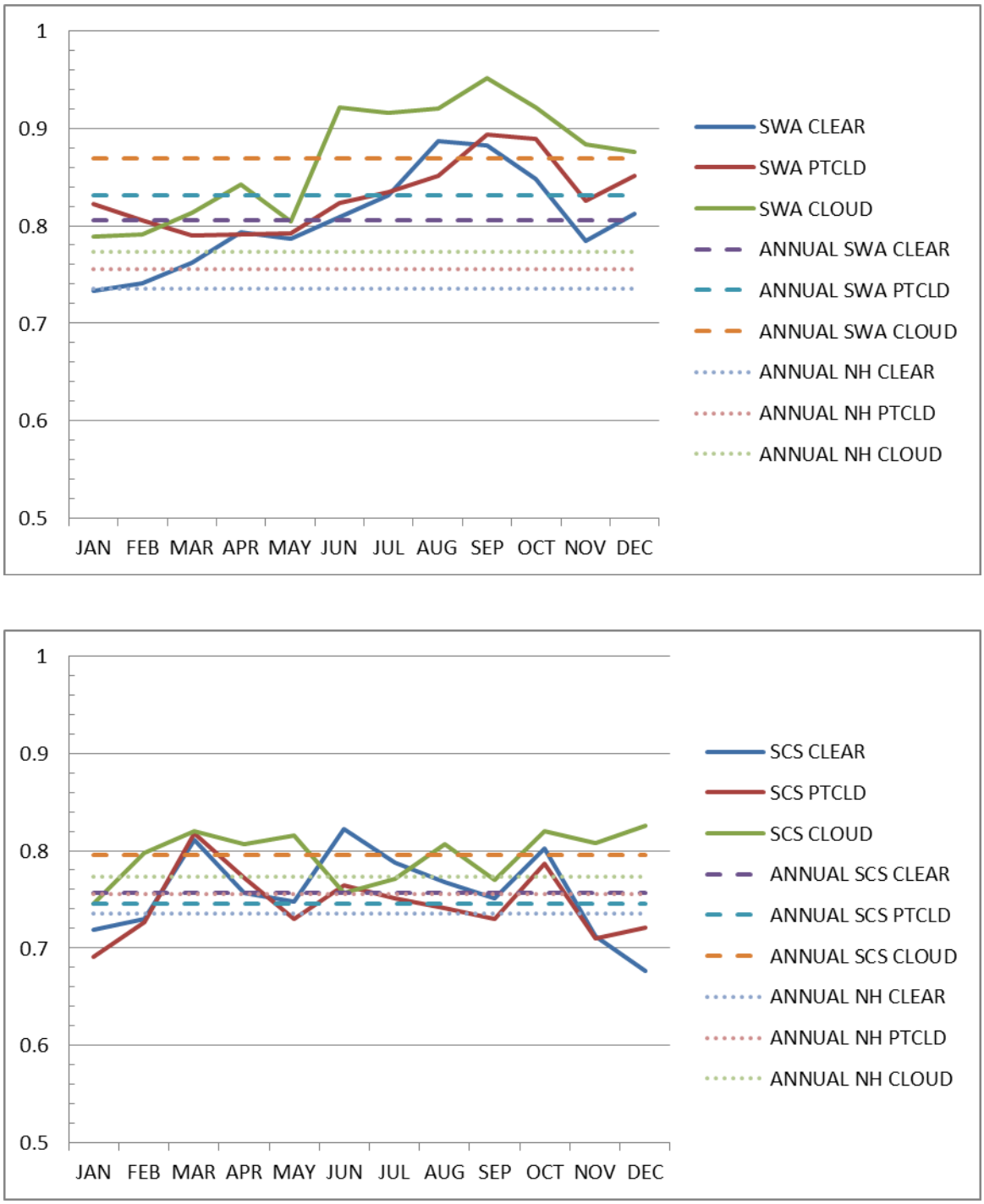


Figure 62. Proportion Correct for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figures 59–61. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

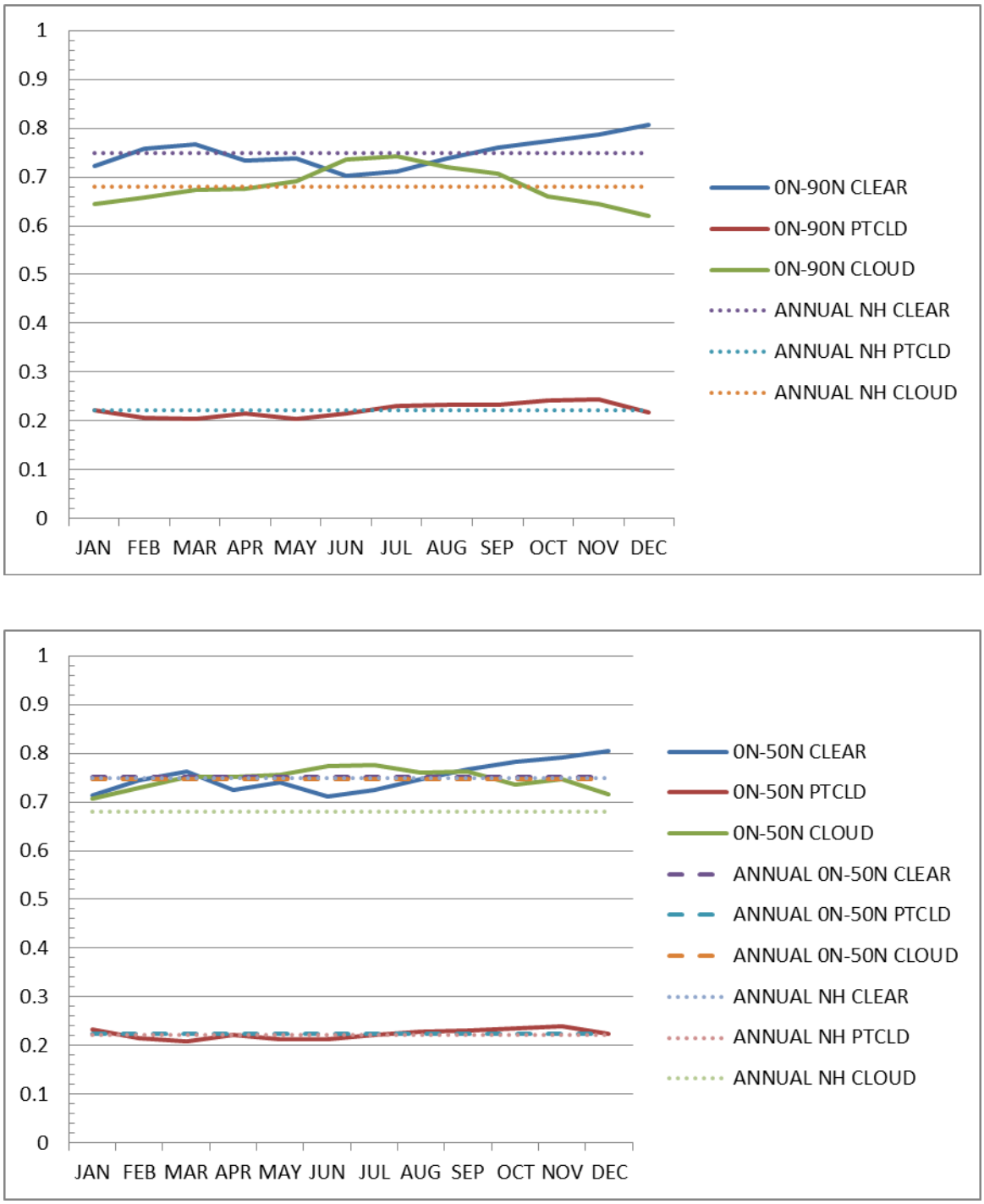


Figure 63. Probability of Detection for each month and all three cloud categories. The top panel is for 0N–90N while the bottom panel is for 0N–50N. This figure should be compared to Figures 64–66. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

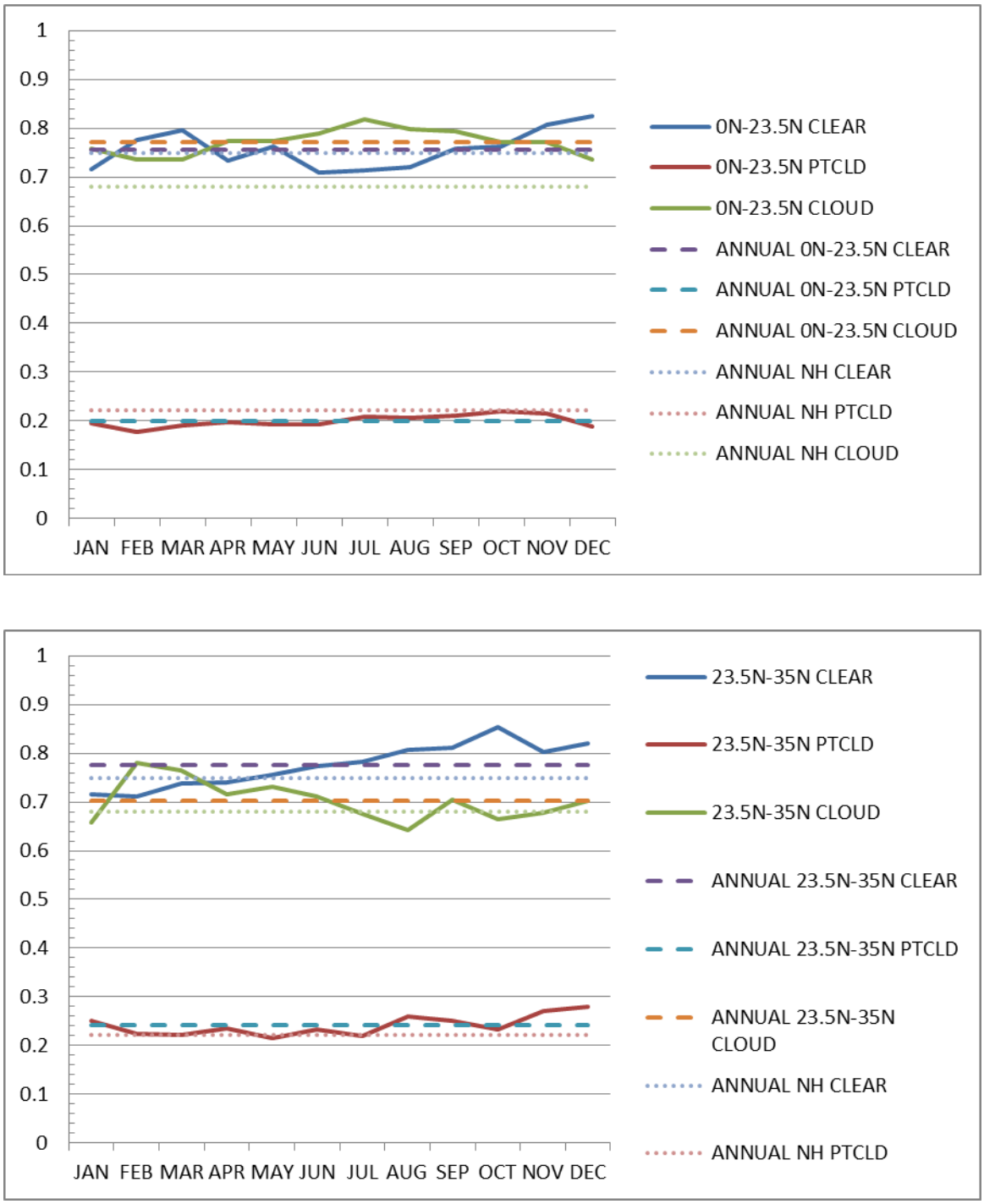


Figure 64. Probability of Detection for each month and all three cloud categories. The top panel is for 0N–23.5N while the bottom panel is for 23.5N–35N. This figure should be compared to Figures 63, 65, and 66. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

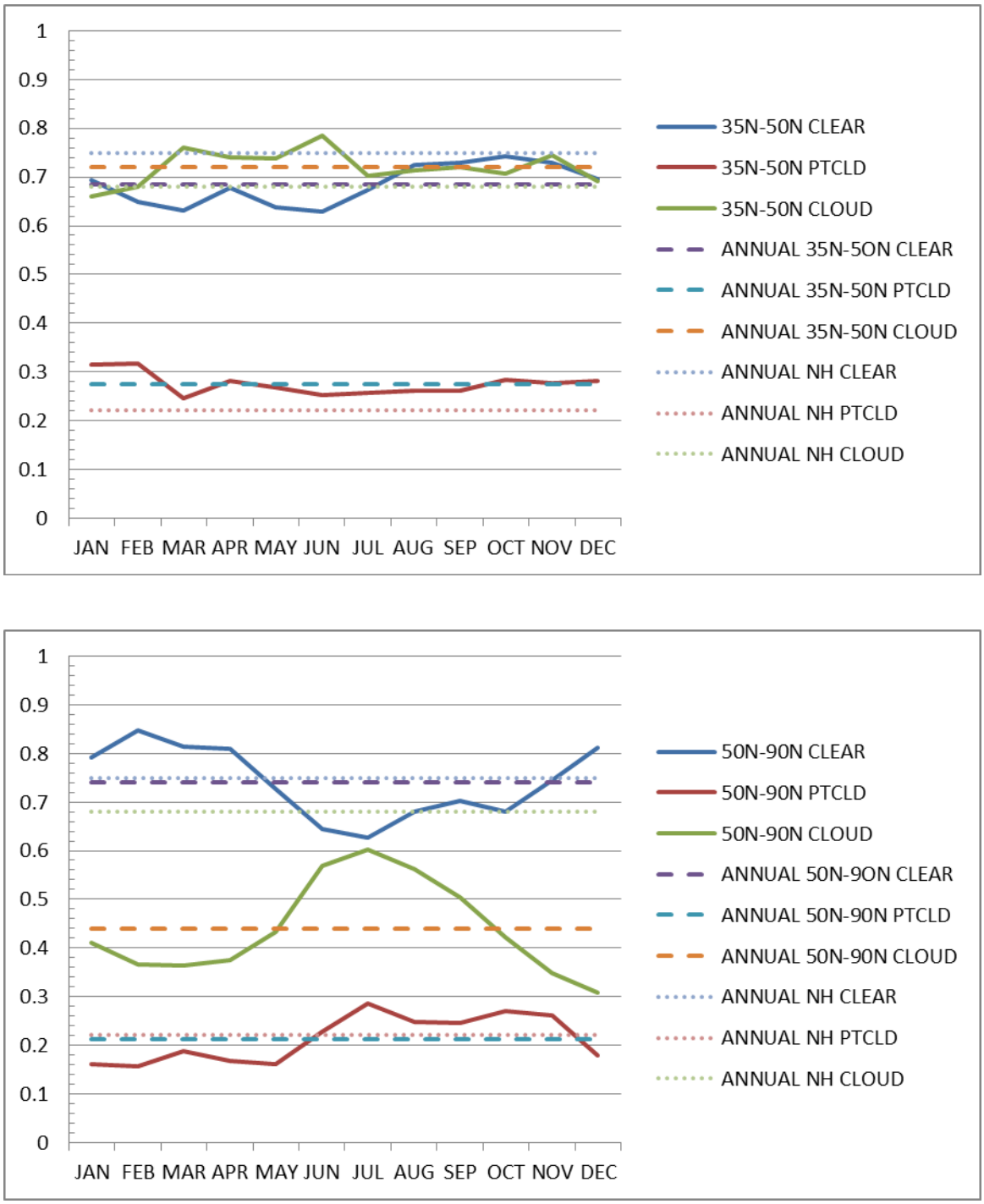


Figure 65. Probability of Detection for each month and all three cloud categories. The top panel is for 35N–50N while the bottom panel is for 50N–90N. This figure should be compared to Figures 63, 64, and 66. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

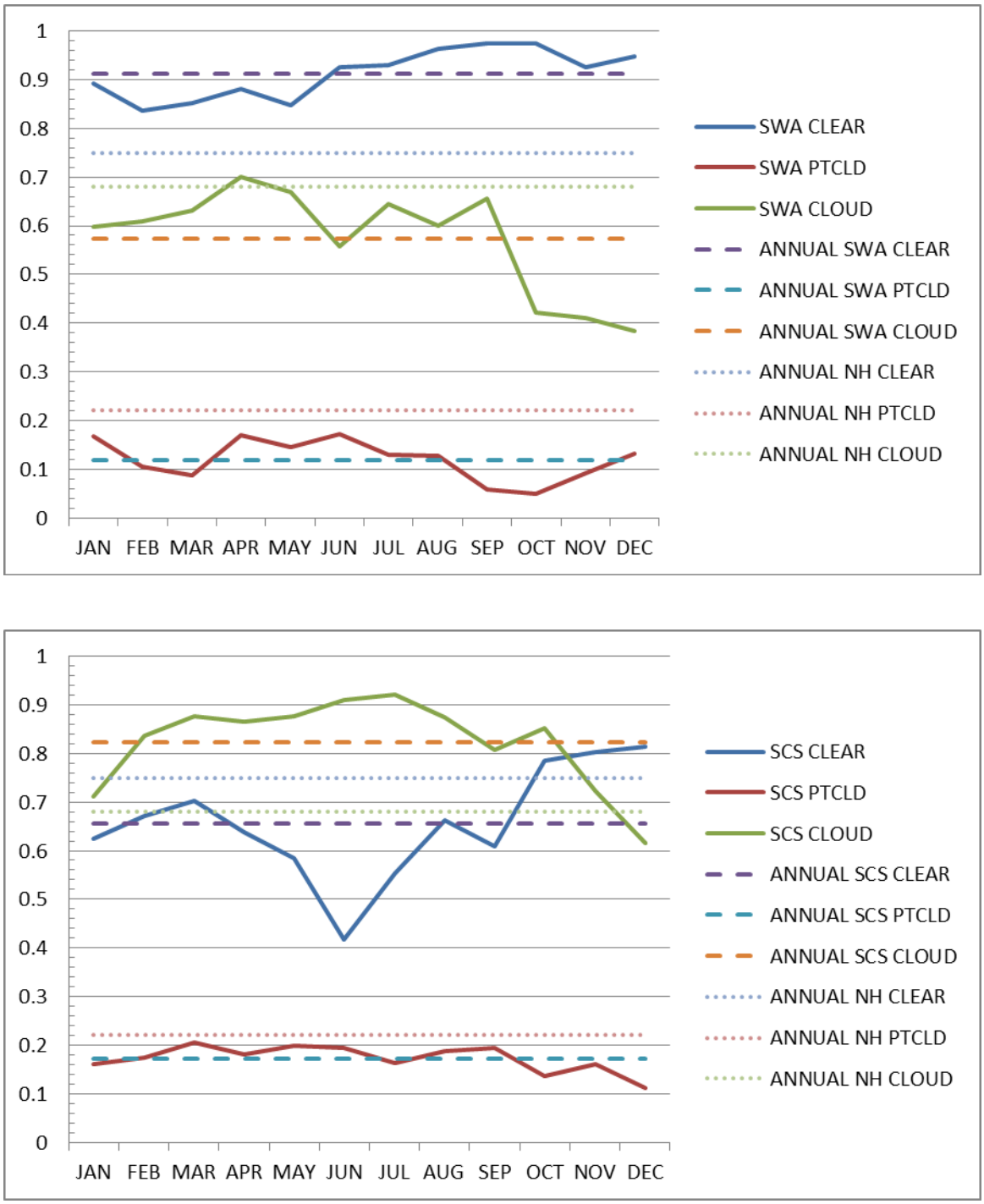


Figure 66. Probability of Detection for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figures 63–65. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

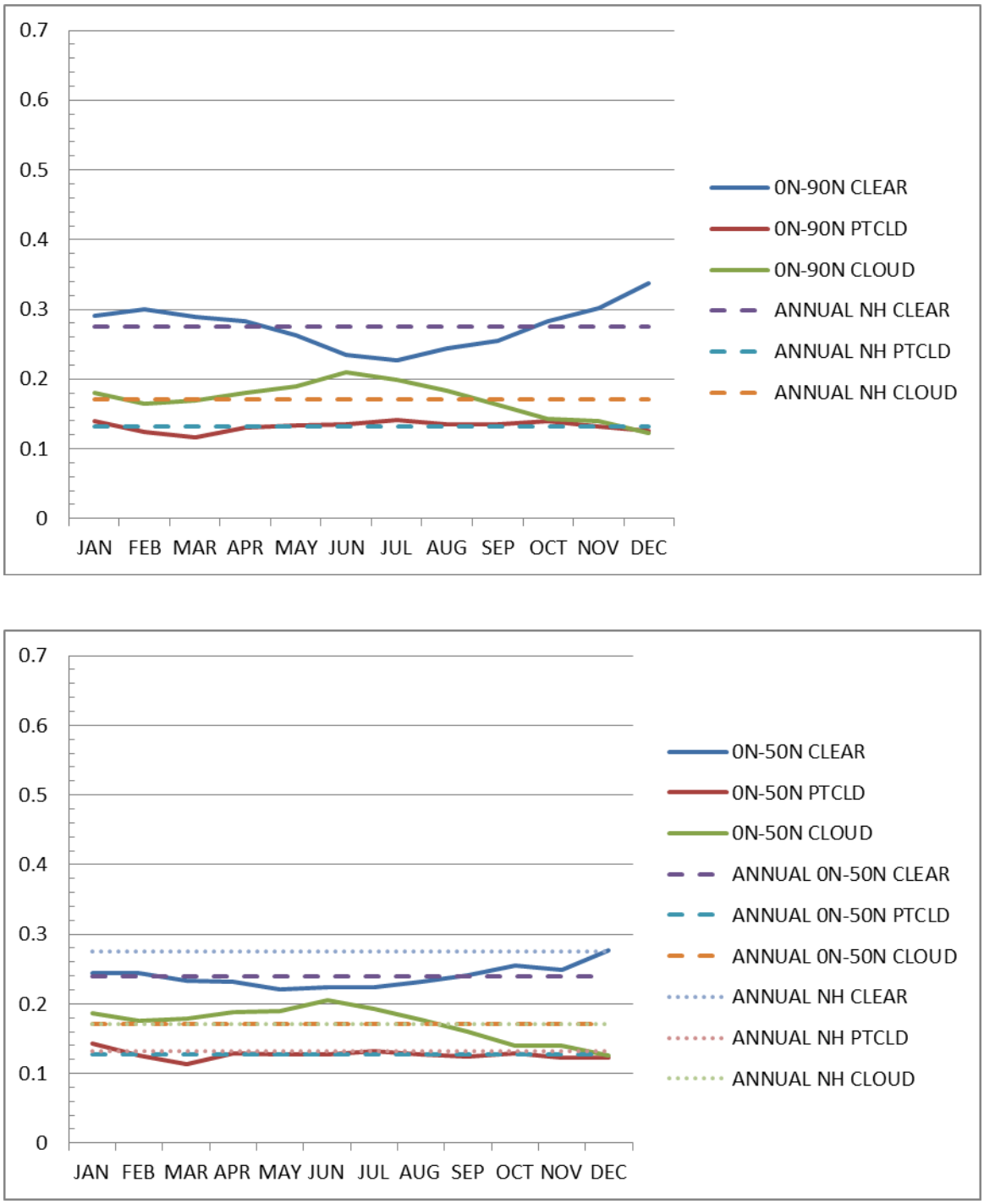


Figure 67. Probability of False Detection for each month and all three cloud categories. The top panel is for 0N–90N while the bottom panel is for 0N–50N. This figure should be compared to Figures 68–70. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

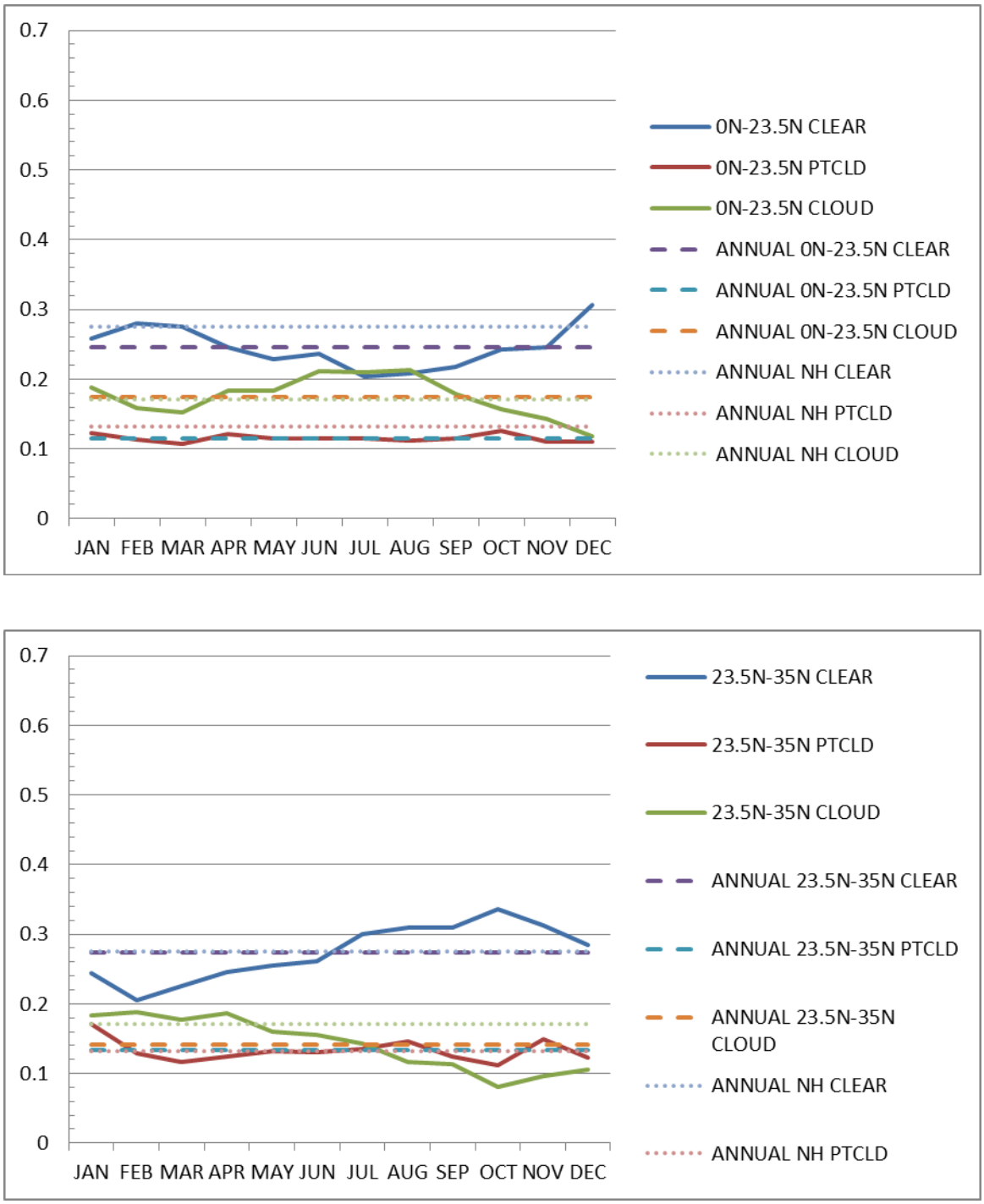


Figure 68. Probability of False Detection for each month and all three cloud categories. The top panel is for 0N–23.5N while the bottom panel is for 23.5N–35N. This figure should be compared to Figures 67, 69, and 70. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

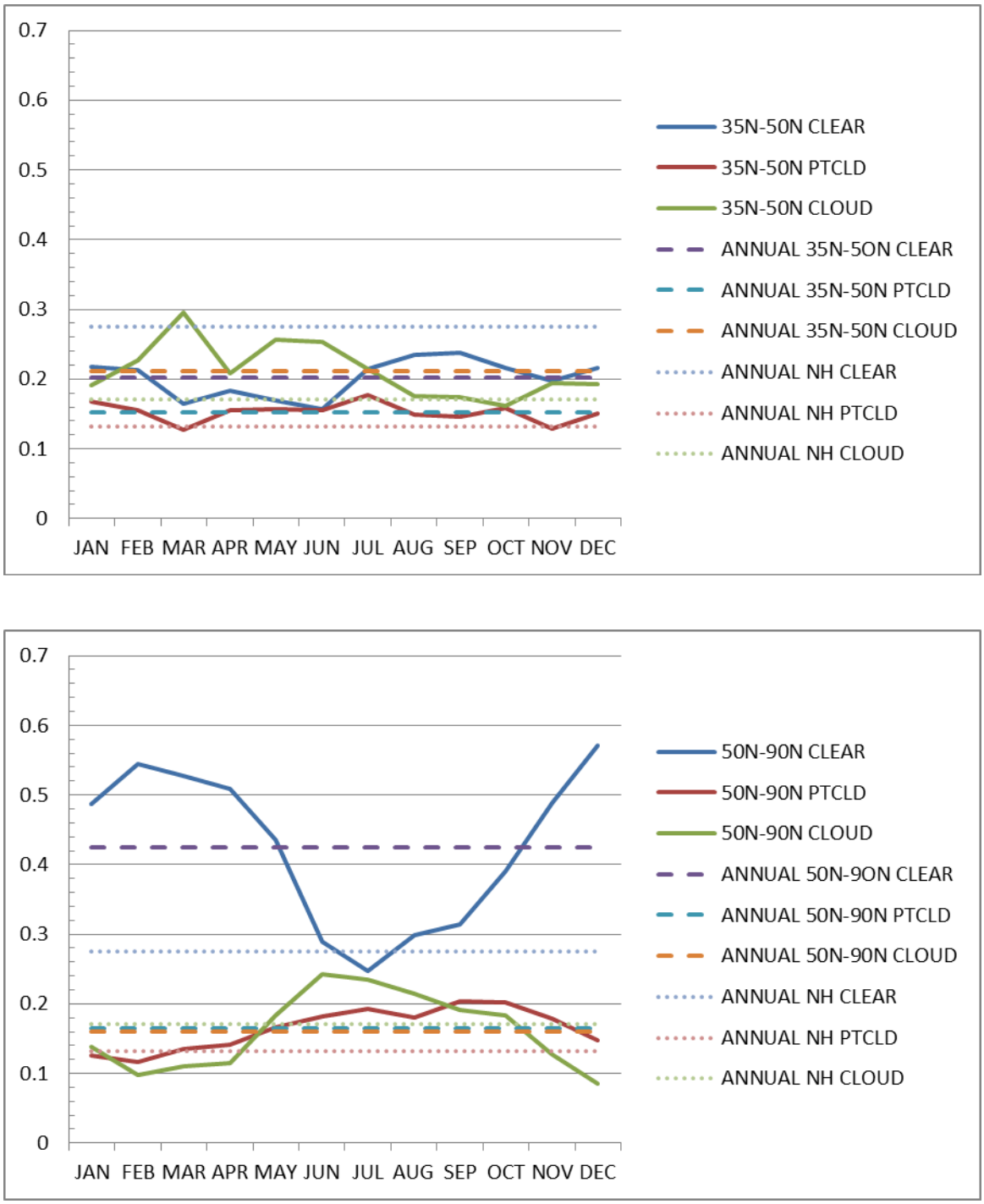


Figure 69. Probability of False Detection for each month and all three cloud categories. The top panel is for 35N–50N while the bottom panel is for 50N–90N. This figure should be compared to Figures 67, 68, and 70. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

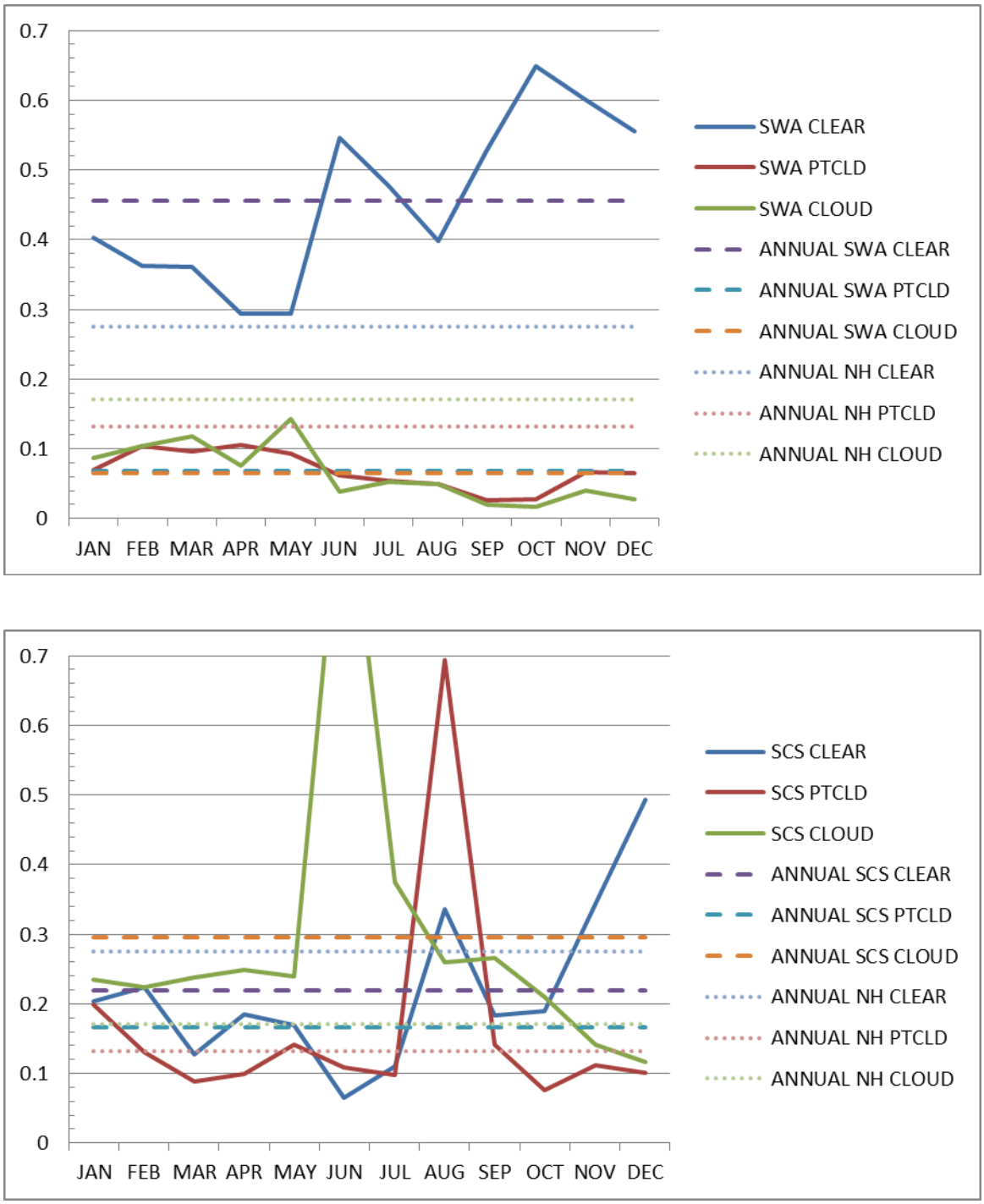


Figure 70. Probability of False Detection for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figures 67–69. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average. The POFD for Clouds in SCS in June was 99%.

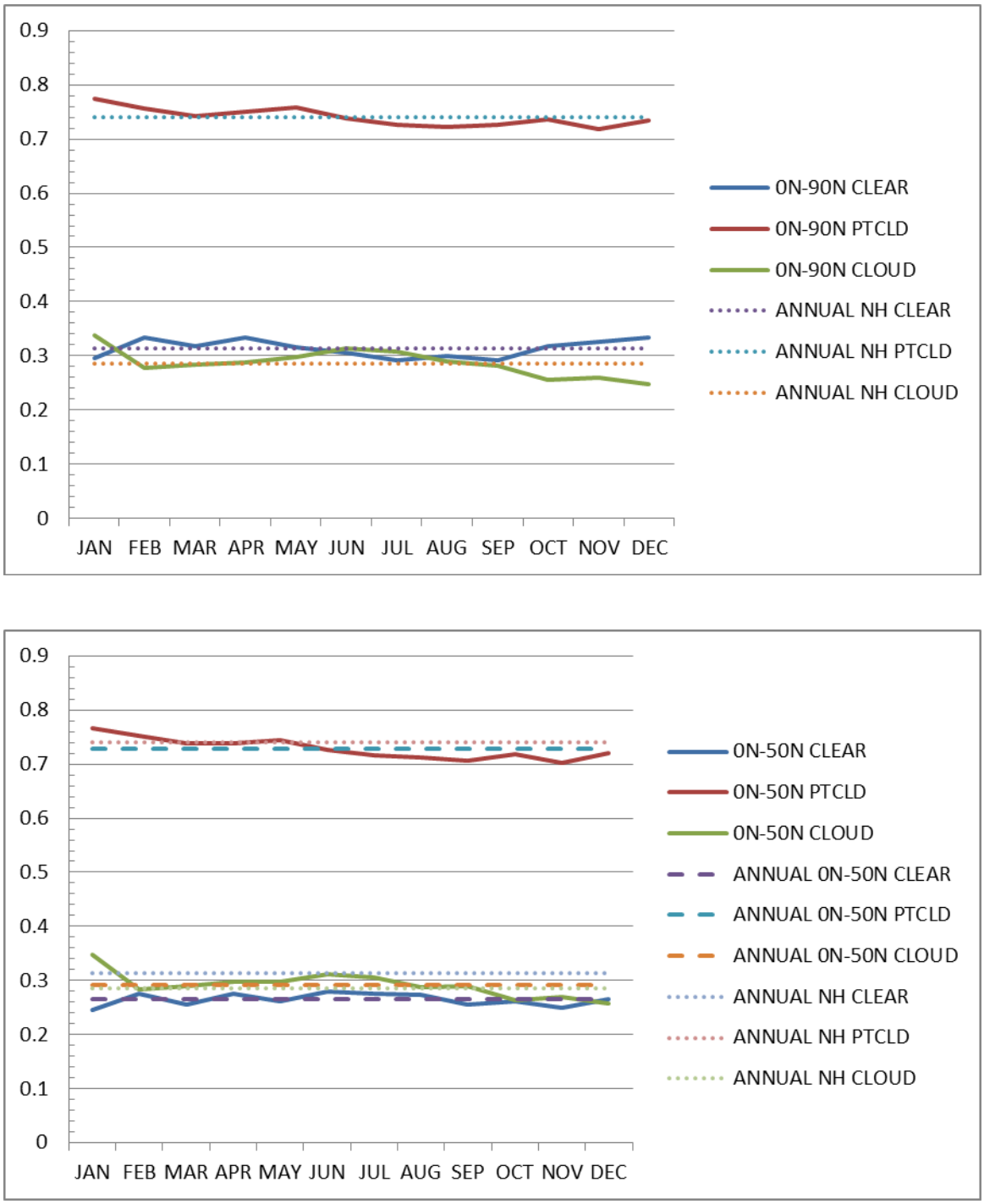


Figure 71. False alarm ratio for each month and all three cloud categories. The top panel is for 0N–90N while the bottom panel is for 0N–50N. This figure should be compared to Figures 72–74. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

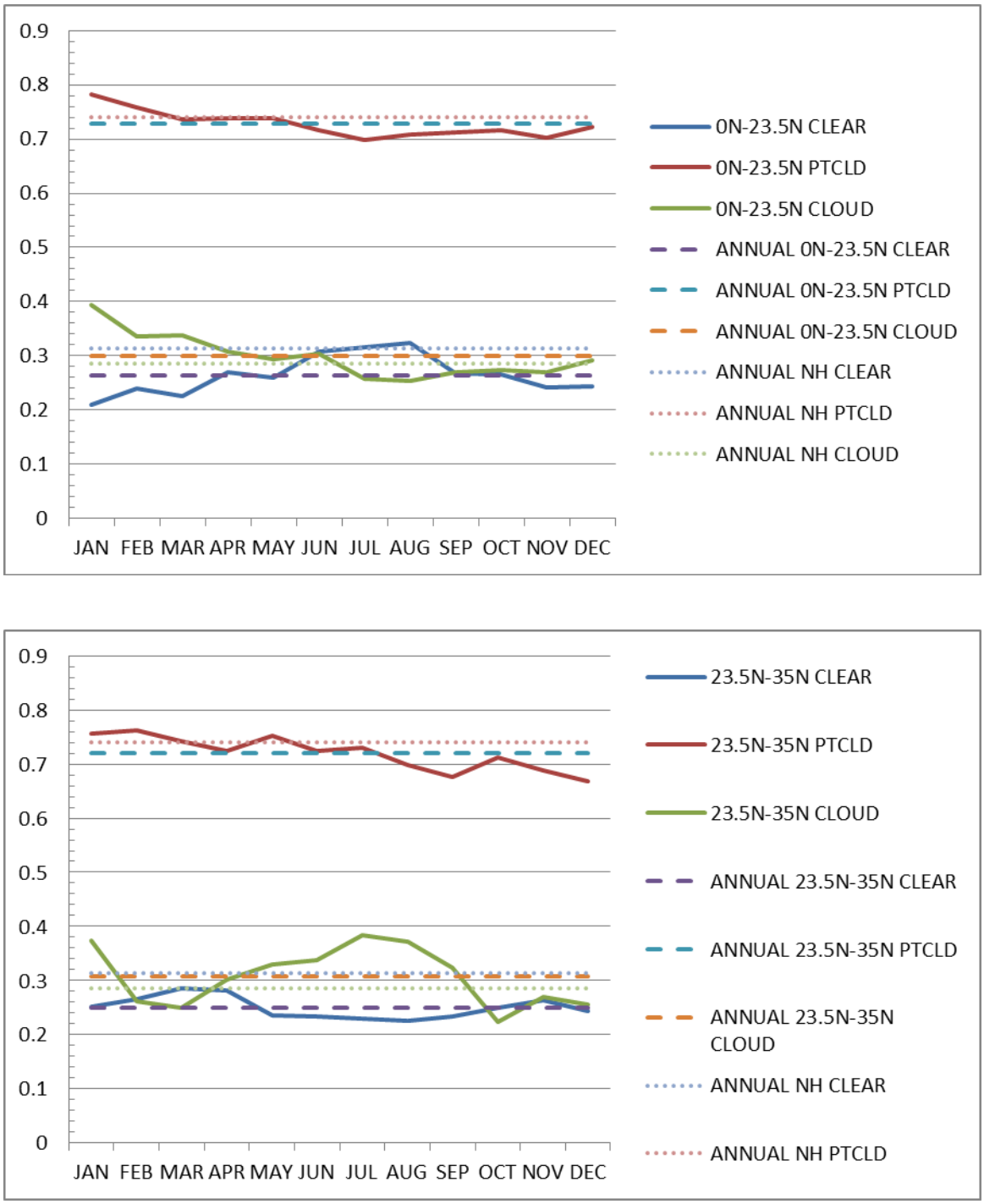


Figure 72. False alarm ratio for each month and all three cloud categories. The top panel is for 0N–23.5N while the bottom panel is for 23.5N–35N. This figure should be compared to Figures 71, 73, and 74. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

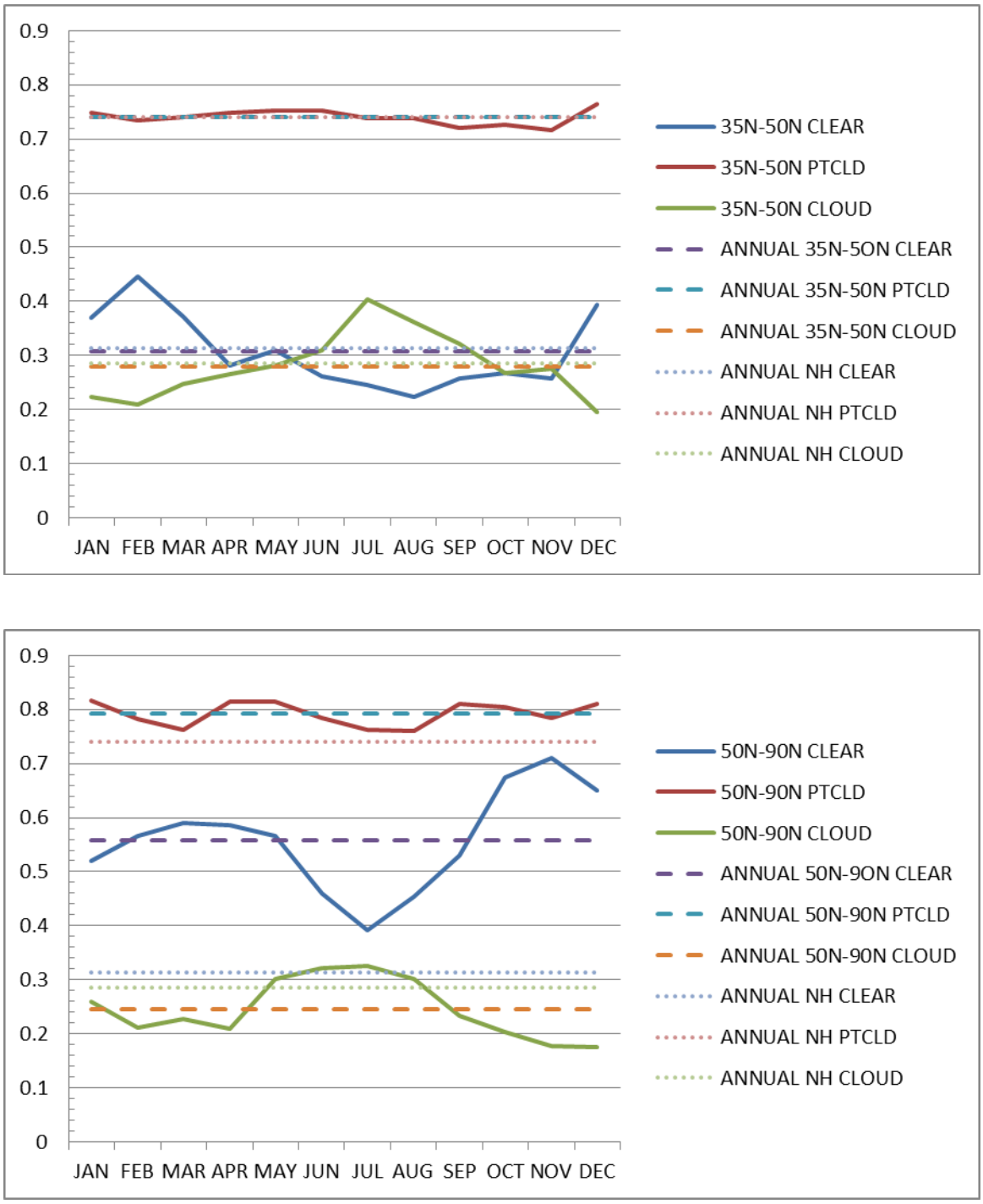


Figure 73. False alarm ratio for each month and all three cloud categories. The top panel is for 35N–50N while the bottom panel is for 50N–90N. This figure should be compared to Figures 71, 72, and 74. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

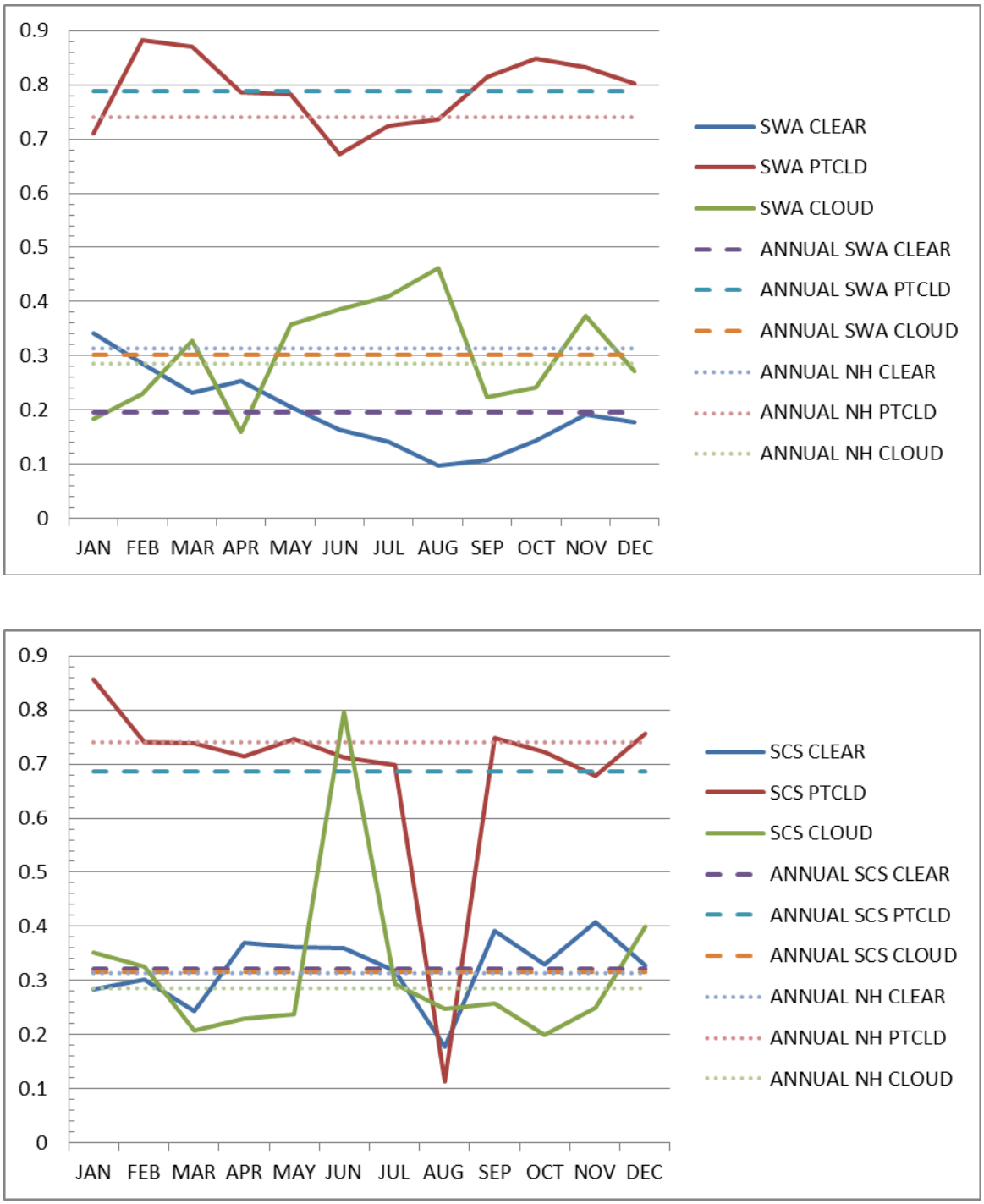


Figure 74. False alarm ratio for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figures 71–73. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

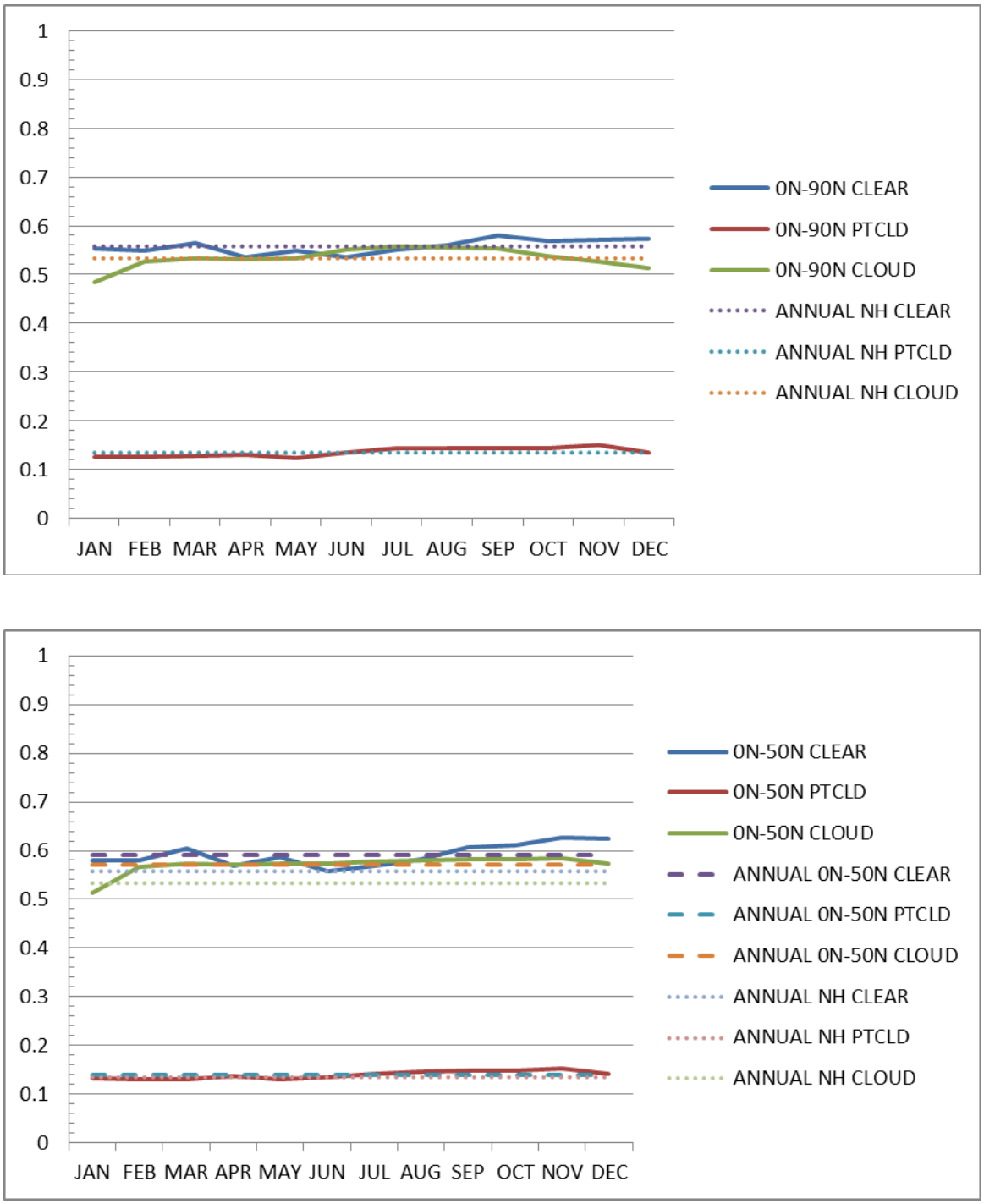


Figure 75. Threat score for each month and all three cloud categories. The top panel is for 0N–90N while the bottom panel is for 0N–50N. This figure should be compared to Figures 76–78. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

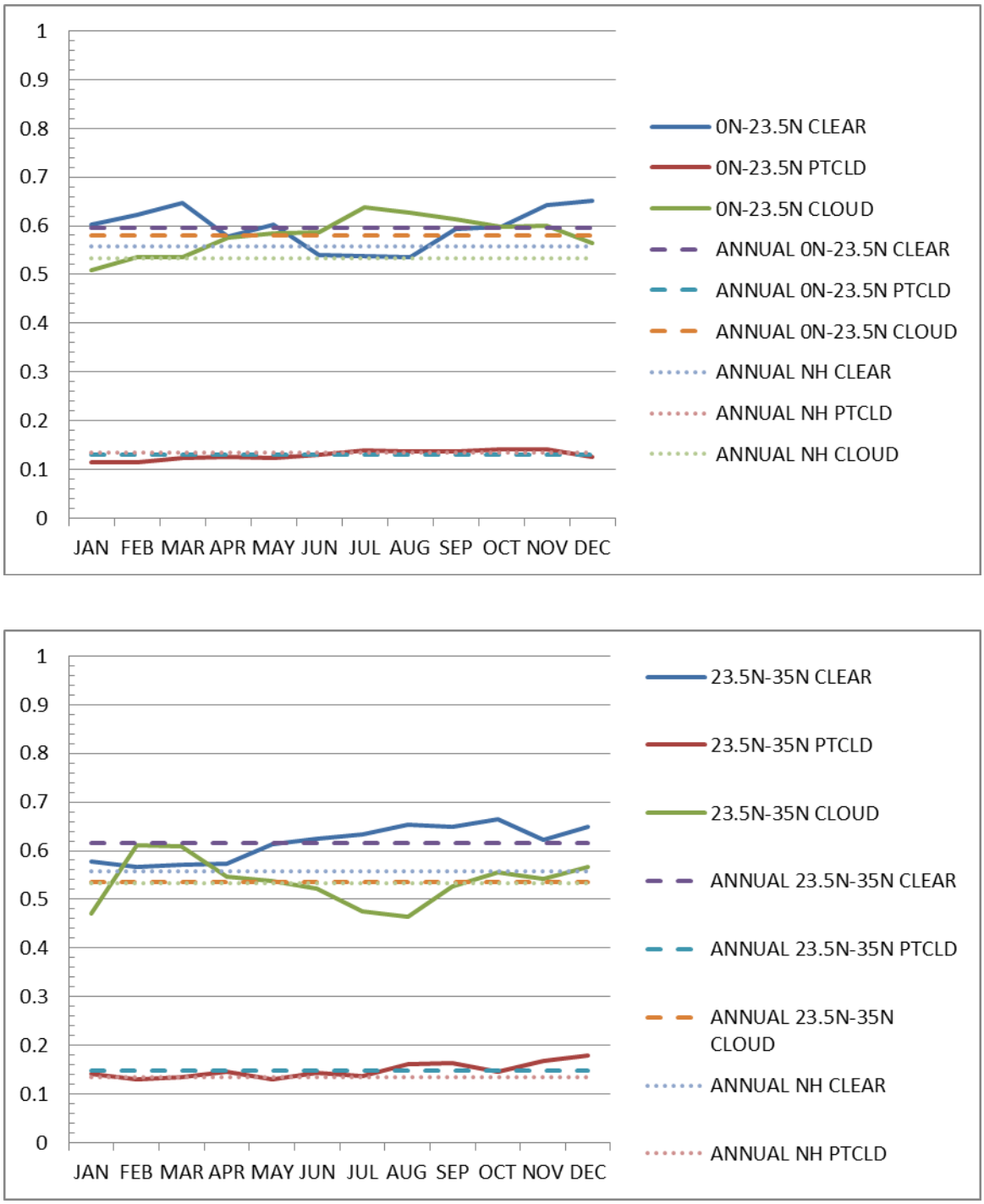


Figure 76. Threat score for each month and all three cloud categories. The top panel is for 0N–23.5N while the bottom panel is for 23.5N–35N. This figure should be compared to Figures 75, 77, and 78. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

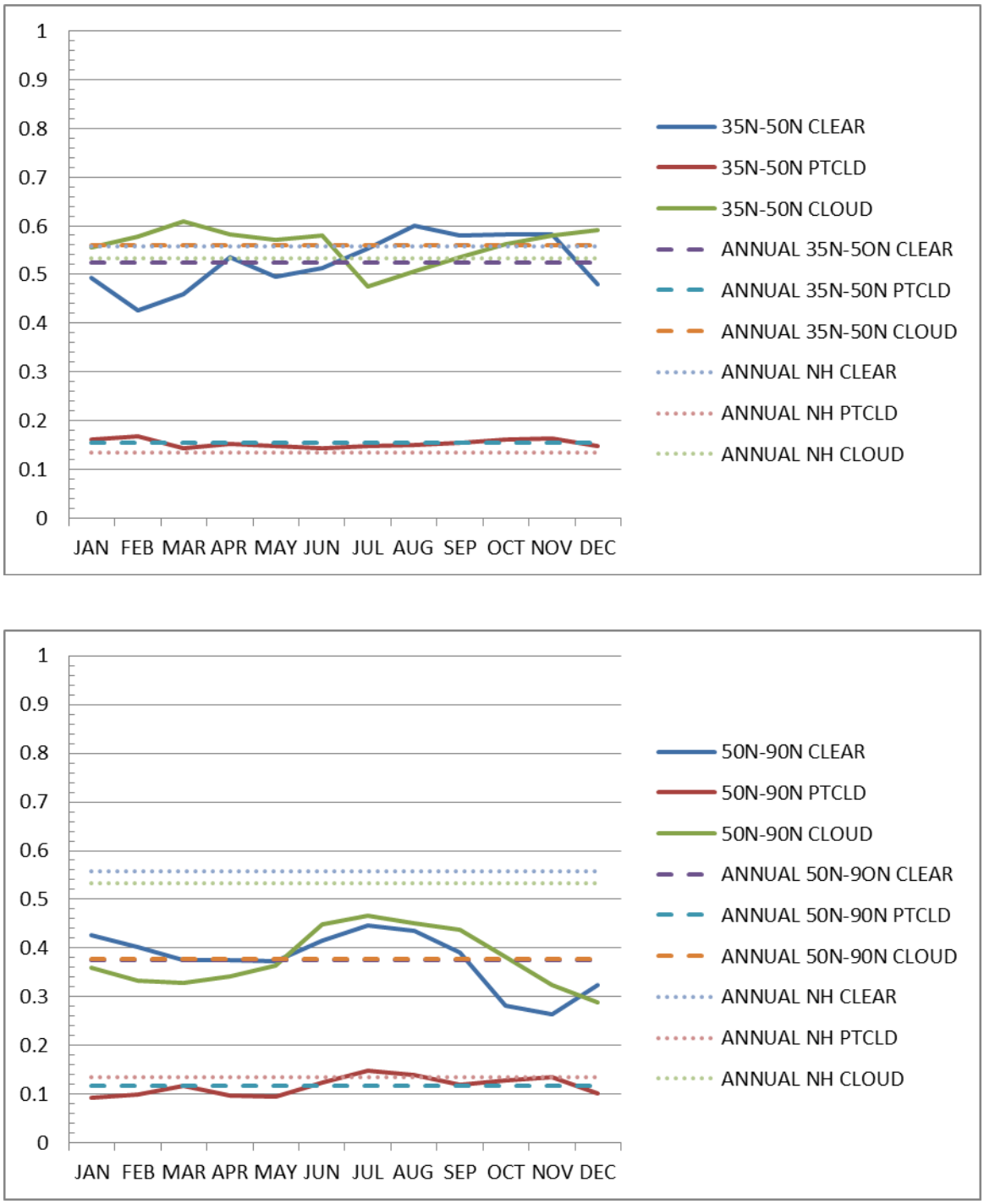


Figure 77. Threat score for each month and all three cloud categories. The top panel is for 35N–50N while the bottom panel is for 50N–90N. This figure should be compared to Figures 75, 76, and 78. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

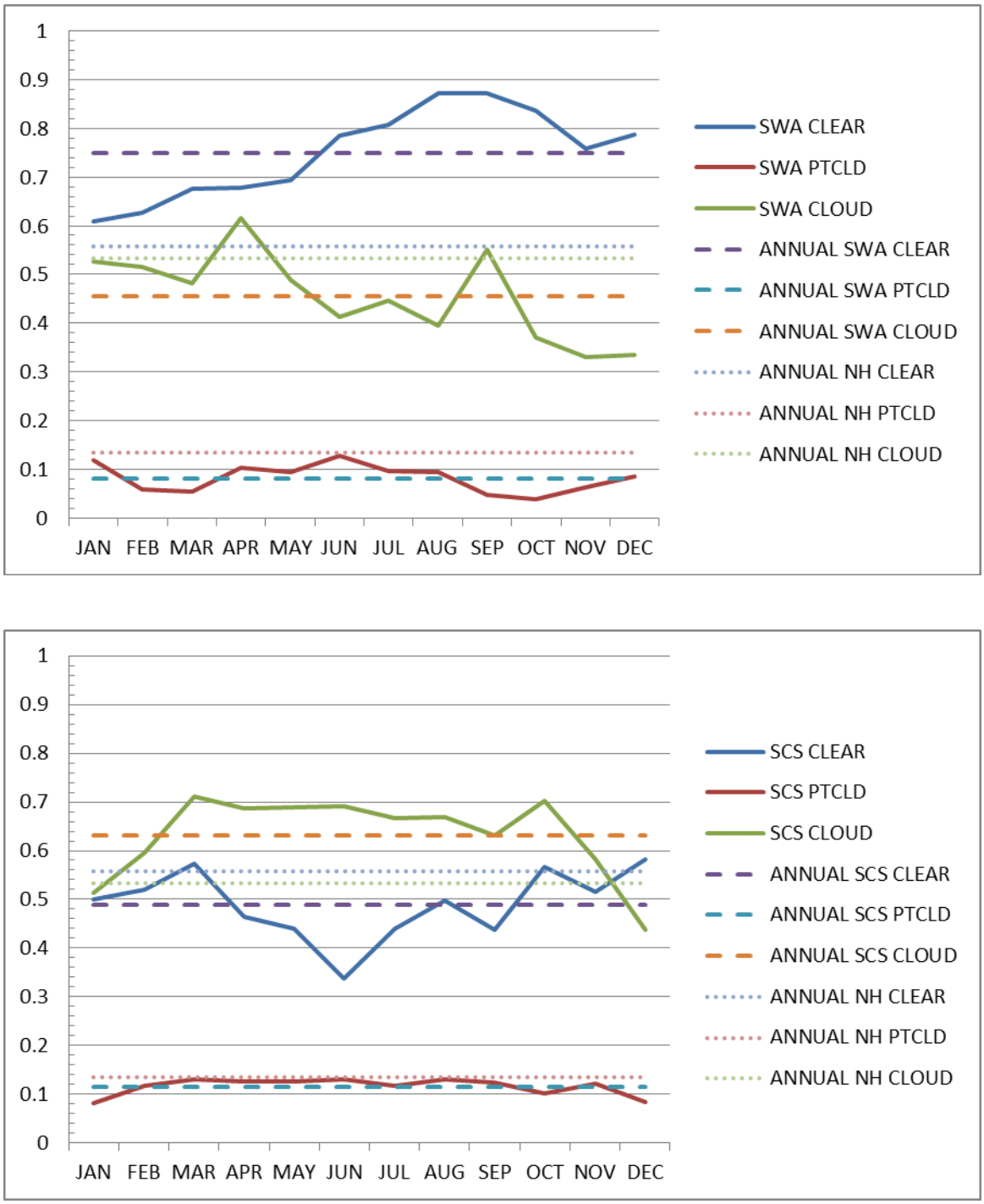


Figure 78. Threat score for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figures 75–77. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

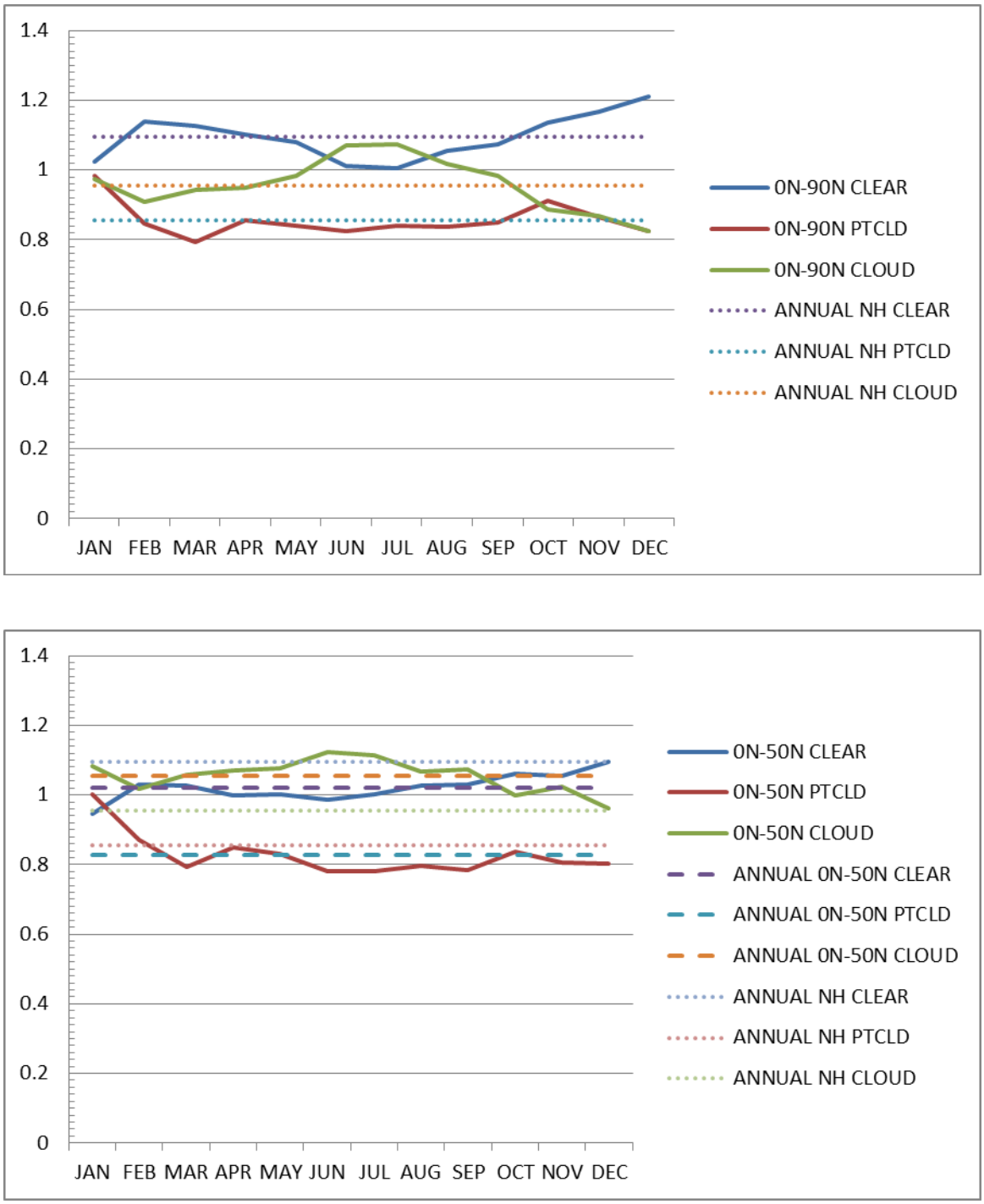


Figure 79. Bias for each month and all three cloud categories. The top panel is for 0N–90N while the bottom panel is for 0N–50N. This figure should be compared to Figures 80–82. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

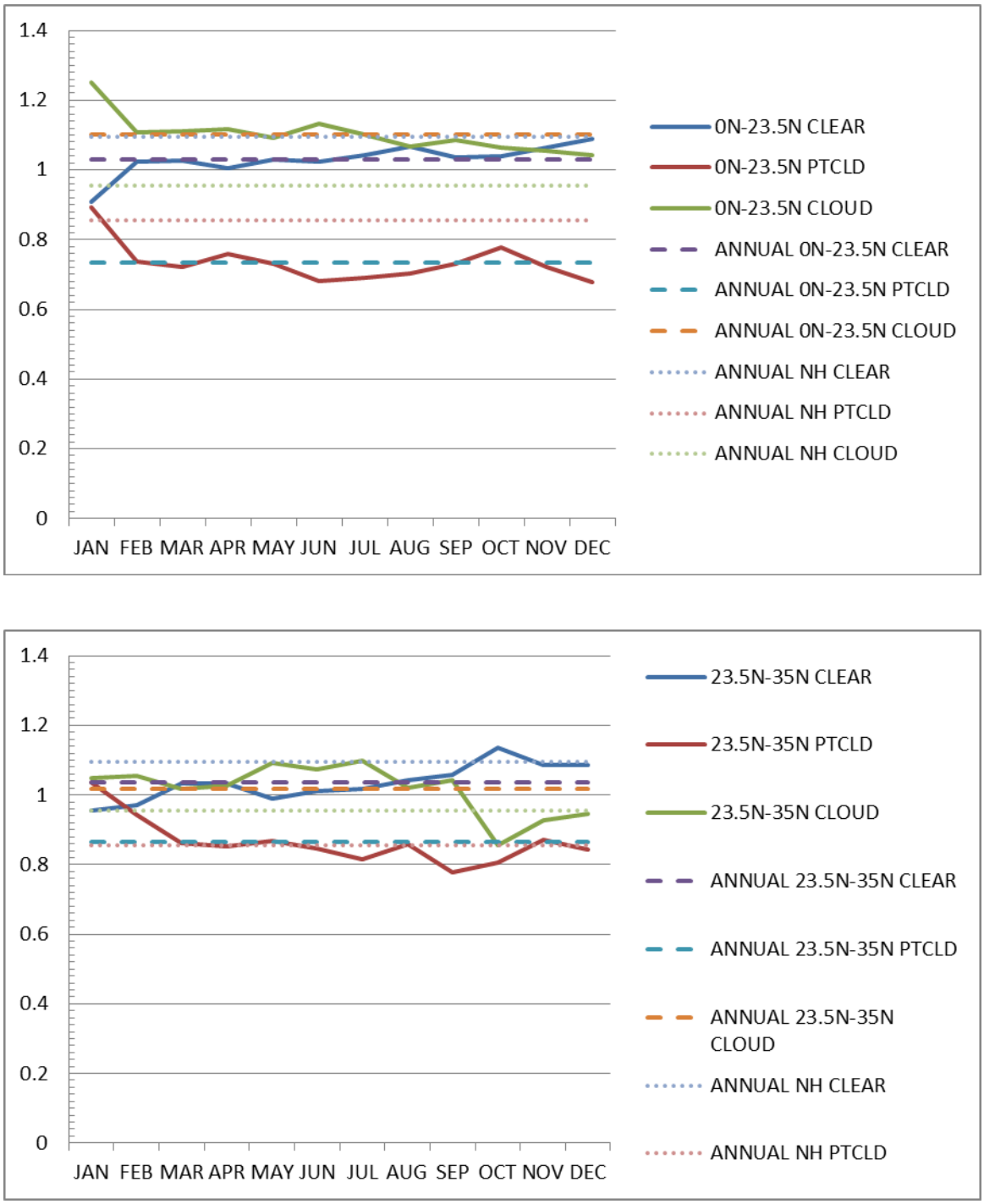


Figure 80. Bias for each month and all three cloud categories. The top panel is for 0N–23.5N while the bottom panel is for 23.5N–35N. This figure should be compared to Figures 79, 81, and 82. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

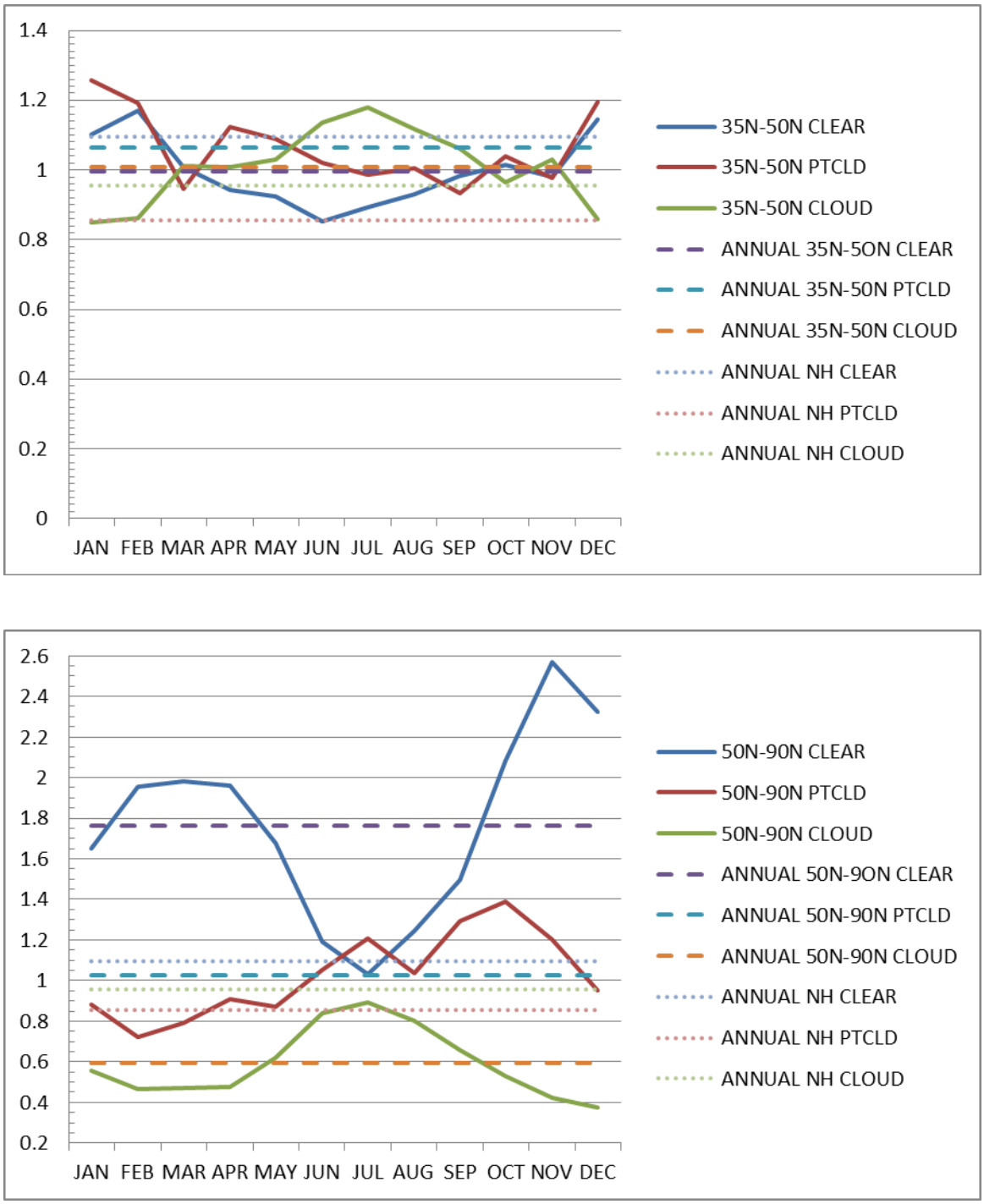


Figure 81. Bias for each month and all three cloud categories. The top panel is for 35N–50N while the bottom panel is for 50N–90N. This figure should be compared to Figures 79, 80, and 82. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

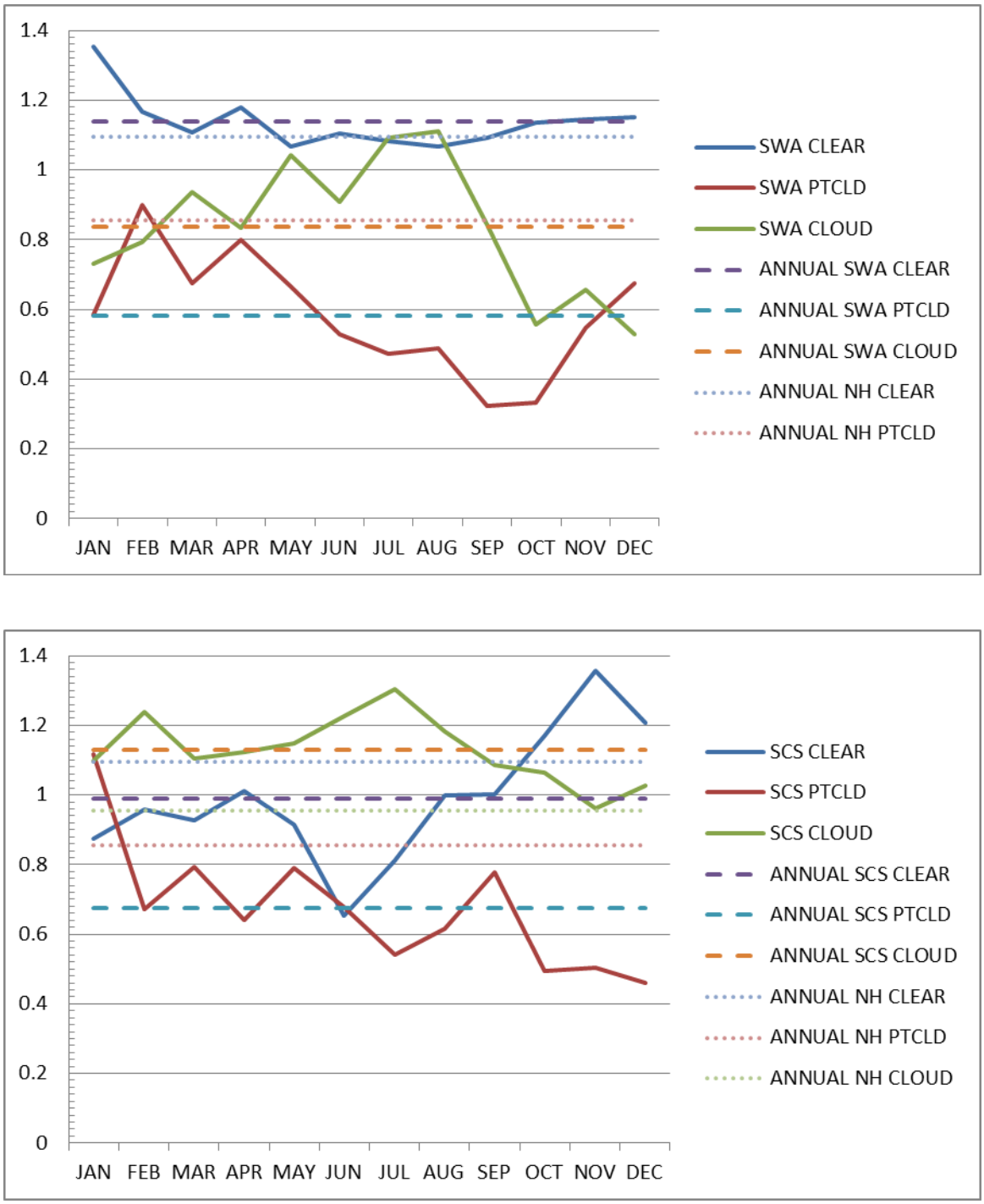


Figure 82. Bias for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figures 79–81. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

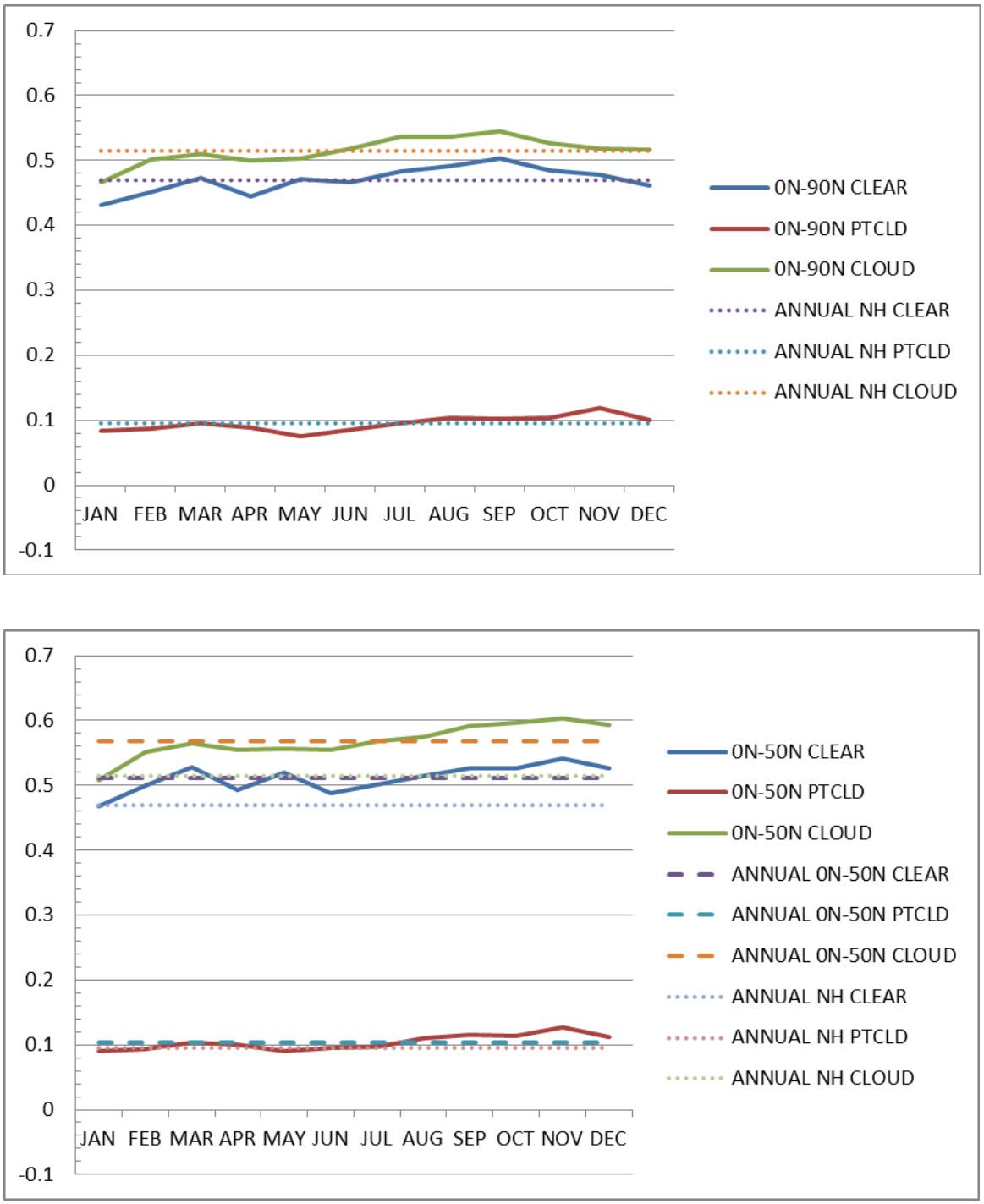


Figure 83. HSS for each month and all three cloud categories. The top panel is for 0N–90N while the bottom panel is for 0N–50N. This figure should be compared to Figures 84–86. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

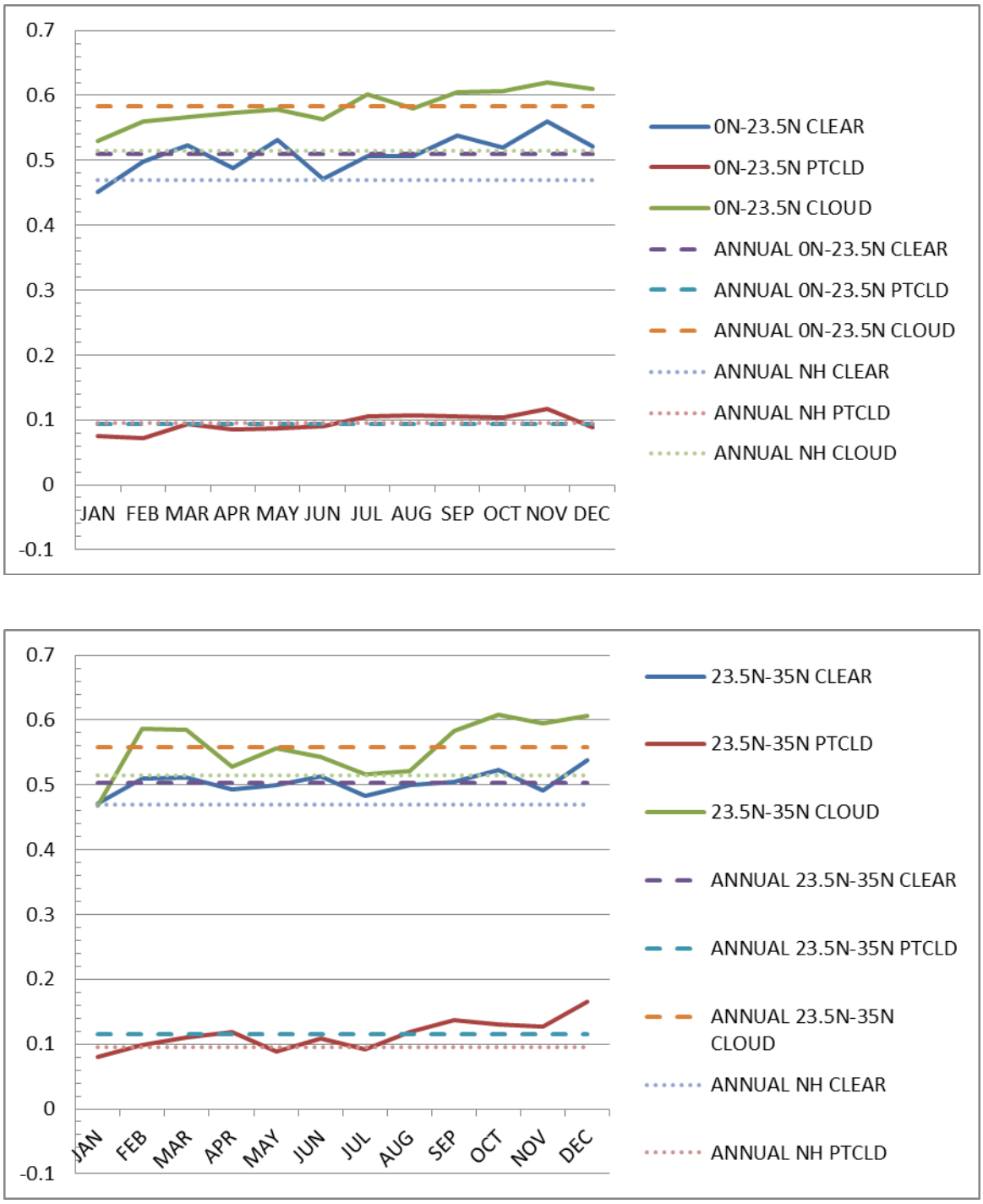


Figure 84. HSS for each month and all three cloud categories. The top panel is for 0N–23.5N while the bottom panel is for 23.5N–35N. This figure should be compared to Figures 83, 85, and 86. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

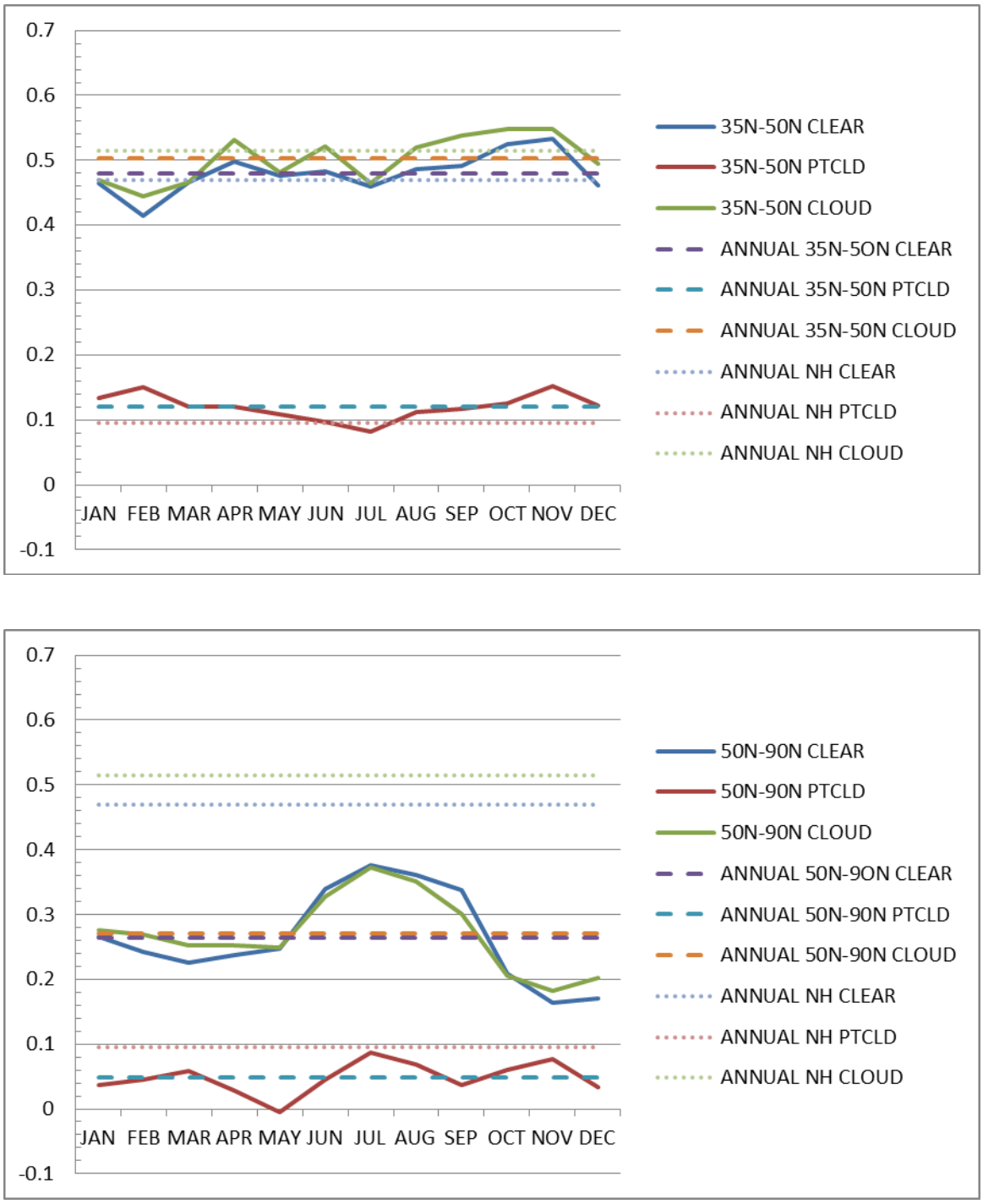


Figure 85. HSS for each month and all three cloud categories. The top panel is for 35N–50N while the bottom panel is for 50N–90N. This figure should be compared to Figures 83, 84, and 86. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

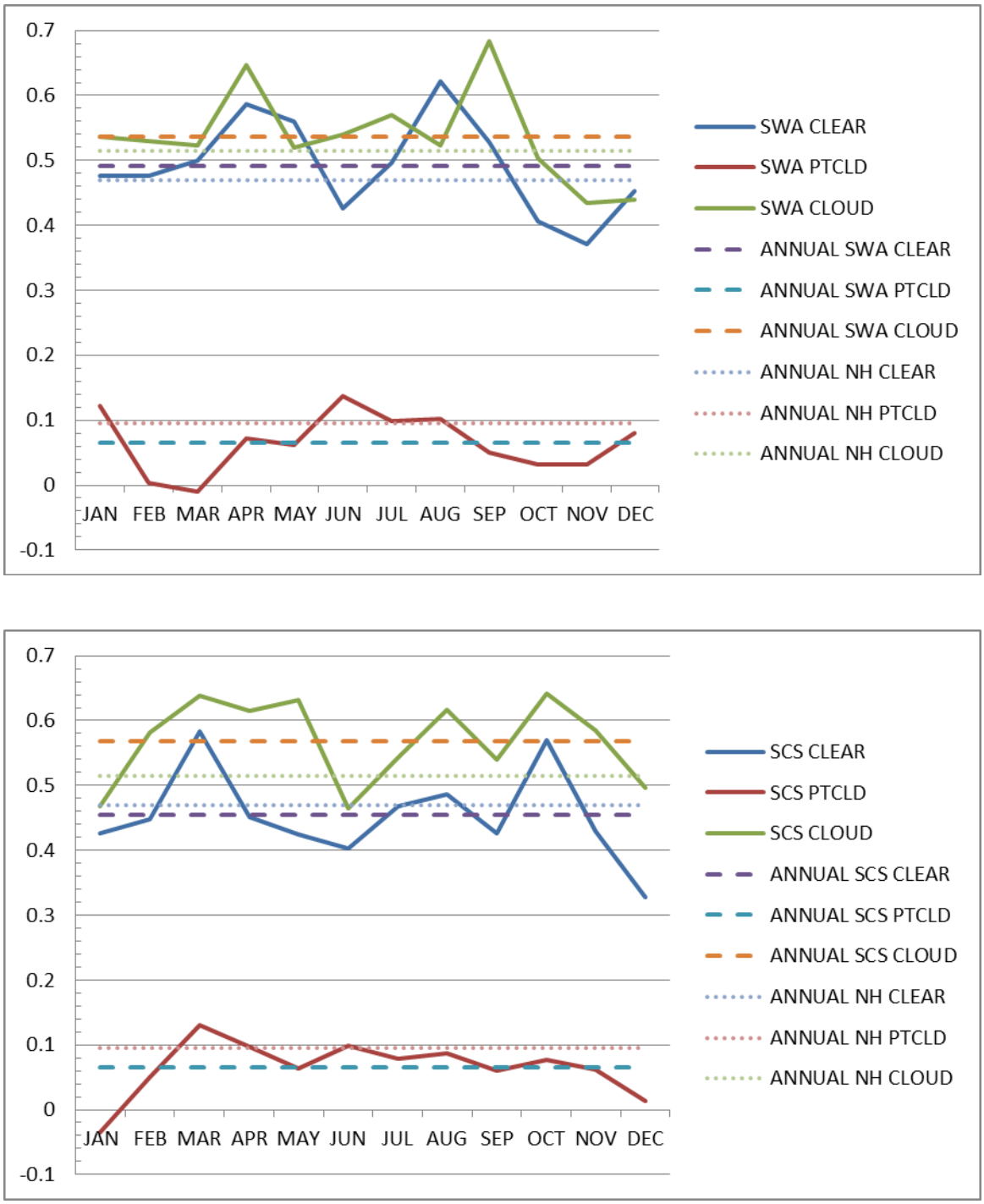


Figure 86. HSS for each month and all three cloud categories. The top panel is for SWA while the bottom panel is for SCS. This figure should be compared to Figures 83–85. Solid lines represent the monthly values for the region, dashed lines represent the annual average for the region, and the dotted line is the Northern Hemisphere annual average.

LIST OF REFERENCES

Bartlett, M.J., 2009: An Evaluation of Methods for Verifying Cloud Distribution in the Worldwide Merged Cloud Analysis. M.S. Thesis. Dept. of Meteorology, Naval Postgraduate School, 68 pp.

Cleary, R. J., 2012a: Verification of Cloud Analyses Used to Support Overhead Imagery Collection. M. S. Thesis. Dept. of Meteorology, Naval Postgraduate School, 134 pp.

Cleary, R. J., 2012b: *Cleary Thesis Brief 09MAR12*, Dept. of Meteorology, Naval Postgraduate School, Monterey, California, 43 pp.

CloudSat Data Processing Center, 2011. [Available at http://www.CloudSat.cira.colostate.edu/dataSpecs.php?prodid=9]

EUMETSAT—Channel Spectral bands, 2011. [Available online at http://www.eumetsat.int/groups/ops/documents/document/pdf_ten_052561_msg1_spctrb nds.pdf]

GOES Imager Channel Notation, 2011. [Available online at http://www.comet.ucar.edu/class/satmet/goeschan.html]

Gustafson, G. B. et al., 2011: Cloud Optical Properties (COP) Integration/Cloud, Dust, Wind Performance Requirements, *Interim Report*, Presented at Air Force Weather Agency, Bellevue, NE, 6 July 2011.

HQ AFWA, 2011: *CDFS II 101*, 16th Weather Squadron, Offutt, Nebraska, 51 pp.

HQ AFWA, 2012: Algorithm Description for the Cloud Depiction and Forecast System II. Offutt AFB, NE 68113.

Hoke, J.E., J.L. Hayes, and L.G. Renniger, 1981 (Rev. March 1985): *Map Projections and Grid Systems for Meteorological Applications*, AFGWC Technical Notes 79/003, USAF AFGWC, Offutt AFB NE 86 pp.

Horsman, S. J., 2007: An Assessment of the World Wide Merged Cloud Analysis Using Interactive Graphics. M.S. Thesis. Dept. of Meteorology, Naval Postgraduate School, 59 pp.

Isaacs, R. G., G. B. Gustafson, T. J. Neu and J. W. Snow, 1994: Improved cloud analysis for CDFS II through the SERCAA research and development program, *Proc. SPIE,* **2309**, Passive Infrared Remote Sensing of Clouds and the Atmosphere II, 15 (December 23, 1994); doi:10.1117/12.196689.

Jarry, J. C., 2005: Analysis of Air Mobility Command Weather Missions Execution Forecasts: Metrics of Forecast Performance and Impacts on War Fighting Operations. M.S. Thesis. Dept. of Meteorology, Naval Postgraduate School, 210 pp.

Mace, G. G., R. Marchand, Q. Zhang, and G. Stephens, 2007: Global hydrometeor occurrence as observed by CloudSat; Initial observations from summer 2006. *Geophysical Research Letters*, **34**, L09808, doi: 10.1029/2006GL029017.

NASA, 2003: NASA Facts—A-Train, 2003: Formation Flying: The Afternoon "A-Train" Satellite Constellation. NASA Facts. [Available online at http://aqua.nasa.gov/doc/pubs/A- Train_Fact_sheet.pdf]

NASA, 2007a: Level 2 GEOPROF Product Process Description and Interface Control Document Algorithm: June 2007, (Version 5.3). [Available online at http://www.CloudSat.cira.colostate.edu /dataICDlist.php?go=list&path=/2B-GEOPROF]

NASA 2007b: CloudSat 2B GEOPROF Quality Statement: May 2007 (Version R04). [Available online at http://www.CloudSat.cira.colostate.edu /dataICDlist.php?go=list&path=/2B-GEOPROF]

NASA 2008: CloudSat Standard Data Products Handbook. [Available online http://www.CloudSat.cira.colostate.edu/CloudSat_documentation/Cloud Sat_Data_Users_H andbook.pdf]

NGIS, 2011: Verification & Validation (V & V) of the World Wide Merged Cloud Analysis (WWMCA): A Review. Technical Note Prepared for the Air Force Weather Agency, Bellevue, NE. SEMSD.1553715944.

Norquist, D. C., 2007: World-Wide Merged Cloud Analysis Verification and Improvement Task 1. Technical Note Prepared for the Air Force Weather Agency, Bellevue, NE.

Ruggiero, F. H., 2000: A Comparison of Cloud-Top Height Estimates from Satellite and NWP Analysis Data. Pre-Prints, *Battlespace Atmospheric and Clouds Impacts to Military Operations (BACIMO) conference*, 24-27 Apr, Fort Collins, CO.

Stephens, G. L., and Coauthors, 2008: CloudSat mission: Performance and early science after the first year of operation, *Journal of Geophyical Research: Atmospheres,* **113,** D00A18, doi:10.1029/2008JD009982.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. International Geophysics Series, Academic Press, 627 pp.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Norm Modlin
   National Reconnaissance Office
   Chantilly, Virginia

4. 16th Weather Squadron (16WS/WXE)
   2d Weather Group (Air Force Weather Agency)
   Offutt AFB, Nebraska

5. 2nd Weather Squadron
   2d Weather Group (Air Force Weather Agency)
   Offutt AFB, Nebraska

6. Col Donald Shannon, USAF
   National System for Geospatial Intelligence
   Springfield, Virginia

7. Tom Murphree, Ph.D.
   Naval Postgraduate School
   Monterey, California

8. Timothy E. Nobis, Ph.D.
   Senior Scientist, SEMS II
   Northrop Grumman Corp
   Offutt AFB, Nebraska

9. Bruce Ford, Clear Science, Inc.
   Jacksonville, Florida