EUMETSAT Satellite Application Facility on Climate Monitoring

Validation Report

CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 2 (CLARA-A2)

Cloud Products

DOI: 10.5676/EUM_SAF_CM/CLARA_AVHRR/V002

Fractional Cloud Cover	**CM-11011**
Joint Cloud property Histogram	**CM-11021**
Cloud Top level	**CM-11031**
Cloud Phase	**CM-11041**
Liquid Water Path	**CM-11051**
Ice Water Path	**CM-11061**

Reference Number:	SAF/CM/DWD/VAL/GAC/CLD
Issue/Revision Index:	2.3
Date:	18.11.2016

Document Signature Table

	Name	Function	Signature	Date
Authors	Karl-Göran Karlsson Joseph Sedlar Abhay Devasthale Martin Stengel Timo Hanschmann Jan Fokke Meirink Nikos Benas Gerd-Jan van Zadelhoff	CM SAF scientist (SMHI) CM SAF scientist (SMHI) CM SAF scientist (SMHI) CM SAF scientist (DWD) CM SAF scientist (DWD) CM SAF scientist (KNMI) CM SAF scientist (KNMI) CM SAF scientist (KNMI)		27/05/2016 19/08/2016
Editor	Rainer Hollmann	Science Coordinator		19/08/2016
Approval	Rainer Hollmann	Science Coordinator		19/08/2016
Release	Martin Werscheck	Project Manager		

Distribution List

Internal Distribution	
Name	**No. Copies**
DWD Archive	1
CM SAF Team	1

External Distribution		
Company	**Name**	**No. Copies**
PUBLIC		1

Document Change Record

Issue/Revision	Date	DCN No.	Changed Pages/Paragraphs
2.0	27/05/2016	SAF/CM/SMHI/VAL/GAC/CLD/2	First version for DRR 2.2
2.1	07/06/2016	SAF/CM/SMHI/VAL/GAC/CLD/2.1	Correction of wrong images in pdf
2.2	19/08/2016	SAF/CM/SMHI/VAL/GAC/CLD/2.2	Final Close_out version with RIDs from DRR 2.2 implemented
2.3	18(06/2016	SAF/CM/SMHI/VAL/GAC/CLD/2.3	Correction for bugfix in CTO global monthly mean and revision of the compliance matrix Update of Figure 6.32, 6.34, 6.36 Update of Tables 1.1, 1.2

Applicable documents

Reference	Title	Code
AD 1	EUMETSAT CM SAF CDOP-2 Product Requirements Document (PRD)	SAF/CM/DWD/PRD, 2.9
AD 2	Requirements Review RR 2.2. document	SAF/CM/CDOP2/DWD/RR22, v1.1

Reference Documents

Reference	Title	Code
RD 1	Product User Manual CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 2 (CLARA-A2) Cloud Products	SAF/CM/DWD/PUM/GAC/CLD, v2.1
RD 2	Algorithm Theoretical Basis Document CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 2 (CLARA-A2) processing chain (level-1 – level-2/2b – level-3).	SAF/CM/DWD/CDOP2/ATBD/CLD, v2.3
RD 3	Algorithm Theoretical Basis Document NWCSAF PPS Cloud Mask Product	NWC/CDOP2/PPS/SMHI/SCI/ATBD/1
RD 4	Algorithm Theoretical Basis Document NWCSAF PPS Cloud Top Product	NWC/CDOP2/PPS/SMHI/SCI/ATBD/3
RD 5	Algorithm Theoretical Basis Document cloud physical products AVHRR	NWC/CDOP2/PPS/SMHI/SCI/ATBD/5
RD 6	Algorithm Theoretical Basis Document CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 2 (CLARA-A2) JCH product	SAF/CM/SMHI/ATBD/JCH, v2.2
RD 7	Algorithm Theoretical Basis Document CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 2 (CLARA-A2) Probabilistic Cloud mask	SAF/CM/SMHI/ATBD/GAC/PBCM, v1.2
RD 8	Validation Report SEVIRI Cloud Products Edition 2 (CLAAS-2)	SAF/CM/KNMI/VAL/SEV/CLD v2.1

TABLE of CONTENTS

1 Executive Summary 14

2 The EUMETSAT SAF on Climate Monitoring 20

3 Introduction to the AVHRR GAC data record 22

4 Cloud products and validation strategy 26

5 Data Sets for Comparison with GAC 32

5.1 SYNOP: manual cloud observations from surface stations 32

5.2 CALIPSO-CALIOP 33

5.3 PATMOS-x 36

5.4 ISCCP 37

5.5 MODIS 37

5.6 DARDAR 38

5.7 UWisc: liquid water path observations from microwave imagers 38

6 Evaluation of CLARA-A2 parameters 40

6.1 Evaluation of AVHRR instantaneous (level-2 and level-2b) products 40
6.1.1 Evaluation against CALIPSO-CALIOP 40
6.1.2 Evaluation against PATMOS-x (level-2b) 56
6.1.3 Evaluation against DARDAR (Cloudsat-CALIPSO) 72

6.2 Evaluation of AVHRR level-3 products (including joint histograms) 77
6.2.1 Evaluation of CLARA-A2 CFC level-3 with SYNOP 77
6.2.2 Evaluation against MODIS 81
6.2.3 Evaluation against ISCCP 85
6.2.4 Evaluation against PATMOS-x 91
6.2.5 Evaluation of CPP products (CPH, LWP, IWP) 96
6.2.6 Joint Cloud property histograms (JCH) 124
6.2.7 Process-oriented comparison of products against other data records 130

6.3 Evaluation of decadal product stabilities 137
6.3.1 Evaluation of decadal stability against SYNOP observations 137
6.3.2 Evaluation of decadal stability against MODIS observations 138
6.3.3 Evaluation of decadal stability against PATMOS-x 139
6.3.4 Decadal stability CPP products 140

7 Conclusions 146

8 References 151

9 Glossary 154

LIST of FIGURES

Figure 3.1 *Local solar times for daytime equator observations for all NOAA satellites from NOAA6 to Metop-B.* 22

Figure 3.2 *Visualisation (same type as in Figure 3.1) of the used satellites in the CLARA-A2 data record.* 24

Figure 5.1 *The Aqua-Train satellites. (Image credit: NASA)* 33

Figure 6.1 *Validation scores Kuipers and Hit Rate as a function of filtered CALIOP-estimated cloud optical depth (see text for explanation). Results are shown for CLARA-A2 in the period 2006-2014 and for CLARA-A1 for a limited validation data record with 99 NOAA-18 orbits 2006-2009.* 42

Figure 6.2 *Global mean CFC (%) from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations. White spots are positions with too limited coverage.* 46

Figure 6.3 *Same as Figure 6.2 but calculated from PPS cloud masks (i.e., CLARA-A2).* 47

Figure 6.4 *Global distribution of the Hit Rate parameter from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations. White spots are positions with too limited coverage.* 48

Figure 6.5 *Global distribution of the Kuipers score from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations. White spots are positions with too limited coverage to allow a confident definition of the Kuipers score.* 48

Figure 6.6 *Global distribution of the CFC Bias from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations. Areas where Bias target requirements are not fulfilled are shown in red (excessive overestimation) and blue (excessive underestimation) colours.* 49

Figure 6.7 *Polar region distribution (Northern Hemisphere to the left and Southern Hemisphere to the right) of the probability of detecting clouds (PODcloudy) from CALIPSO-CALIOP calculated from* ***Polar night collocations*** *for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations.* 49

Figure 6.8 *Polar region distribution (Northern Hemisphere to the left and Southern Hemisphere to the right) of the daytime (Polar Summer) probability of detecting clouds (PODcloudy) from CALIPSO-CALIOP calculated from* ***Polar day collocations*** *for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations.* 50

Figure 6.9 *Mean cloud top height (CTH) deviations from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Results are given as a function of filtered cloud optical depths (see text for details).* 53

Figure 6.10 *a) Global mean cloud fraction [%] for PATMOS-x (green) and CLARA-A2 (red). Daily averages are computed from all (ascending + descending) satellite overpasses. The R-value (correlation coefficient) is provided in the lower right. Global averages are area-weighted. b) Daily global mean cloud fraction [%] difference, defined as CLARA-A2 – PATMOS-x. The mean bias error [%] and bias-corrected RMSE [%] in globally-averaged daily cloud fraction is provided in the upper right.* 57

Figure 6.11 *2D relative frequency histograms for CLARA-A2 vs. PATMOS-x daily-averaged global cloud fraction from 1982-2014. Dashed lines show deviations from the 1:1 line of +5 % and -5 %.* 58

Figure 6.12 *Same as in Figure 6.10, but with global, daily-averaged cloud fraction for a) afternoon (PM) and b) overnight (AM) local satellite overpass times. c) Daily difference in global average cloud fraction for PM (black) and AM (gray) overpasses.* 59

Figure 6.13 *a) Global mean cloud top pressure [hPa] for PATMOS-x (green) and CLARA-A2 (red). Daily averages are computed from all (ascending + descending) satellite overpasses. The R-value (correlation coefficient) is provided in the lower right. Global averages are area-weighted. b) Same as a) but only for pixels both having clouds (common cloud mask)c) Daily global mean cloud top pressure [hPa] difference, defined as CLARA-A2 – PATMOS-x. Results are given both for all cases (black) and for the common cloud mask (grey). The mean bias error, bias-corrected RMS error (both in hPa) and correlation in globally-averaged daily cloud top pressure are provided in the upper right in figures a) and b).* 62

Figure 6.14 *Relative frequency distributions [%] of all valid afternoon (PM, full lines) and overnight (AM, dashed lines) overpasses of level-2b CTP values for CLARA-A2 (red) and PATMOS-x (green) during 1982-2014 for a) January and b) July.* 64

Figure 6.15 *Mean daytime cloud top height in km from NOAA-18 for July 2008. Left: CLARA-A2; right: PATMOS-x. Grey areas indicate no data because no clouds were detected or the solar zenith angle was too high during the entire month* 65

Figure 6.16 *Pixel-level comparison between CLARA-A2 and PATMOS-x daytime CTH from NOAA-18 for July 2008. Left: scatter-density plot in which the colours indicate the number of pixels (level-2b grid cells) with the particular CLARA and PATMOS CTH values; right: 1-dimensional histograms.* 66

Figure 6.17 *Mean liquid cloud optical thickness (top) and effective radius from NOAA-18 for July 2008. Left: CLARA-A2; right: PATMOS-x. Grey areas indicate no data because no clouds were detected or the solar zenith angle was too high during the entire month.* 67

Figure 6.18 *Pixel-level comparison between CLARA-A2 and PATMOS-x liquid COT (top: note that the logarithm of COT is shown), liquid REFF (middle), and LWP (bottom) from NOAA-18 for July 2008. Left: scatter-density plots in which the colours indicate the number of pixels (level-2b grid cells) with the particular CLARA and PATMOS parameter values; right: 1-dimensional histograms.* 68

Figure 6.19 *As Figure 6.17, but now for NOAA-17, and only liquid REFF is shown.* 69

Figure 6.20 *As Figure 6.18, but now for NOAA-17, and only liquid REFF is shown.* 69

Figure 6.21 *As Figure 6.17, but now for ice cloud properties.* 70

Figure 6.22 *As Figure 6.21, but now for ice cloud properties.* 71

Figure 6.23 *As Figure 6.21, but now for NOAA-17, and only ice REFF is shown.* 72

Figure 6.24 *As Figure 6.22, but now for NOAA-17, and only ice REFF is shown.* 72

Figure 6.25 *CLARA-A2 cloud phase hit rate, probability of detections and false alarm ratios for both liquid and ice clouds as a function of the integrated optical thickness from the top of the cloud. Results are for NOAA-18 collocations in January 2008. Only single cloud phase DARDAR columns were used in these statistics and both day and night observations were taken into account.* 73

Figure 6.26 *Ice cloud optical thickness distribution comparing the DARDAR and CLARA-2 retrieved collocated values. The blue dashed line shows the 1-1 line with the greyscales indicating the regions enclosing the 20, 40, 60, 75, and 90% of points with the highest occurrence frequency.* 74

Figure 6.27 *Comparison of CLARA-A2 ice effective radius and DARDAR weighted effective radius from cloud top to an optical depth of 1 (or to cloud base if the total optical depth is smaller than 1). The left plot shows 1D-histograms with CLARA-A2 indicated in red and DARDAR in black; on the right a scatter density plot is shown. The dynamic range of the DARDAR retrievals is a lot larger resulting in no correlation between the two distributions. The greyscales indicate regions enclosing the 10, 30, 50, 70, and 90% of points with the highest occurrence frequency.* 75

Figure 6.28 *Left panel: CLARA-A2 IWP vs. DARDAR IWP. The yellow line depicts the median and orange the 16th/84th percentiles of the CLARA-A2 distribution at the local DARDAR IWP. Right panel: 1D-histograms of DARDAR and CLARA-A2 IWP for the same collocations. The greyscales indicate regions enclosing the 10, 20, 40, 60, and 75% of points with the highest occurrence frequency.* 75

Figure 6.29 *Time series of mean cloud cover for CLARA-A2 (red), CLARA-A1 (grey), and SYNOP (black) (upper panel), bias-corrected RMSE (second panel), bias (third panel), and the number of stations (lower panel) normalized to 1 for the entire period 1982-2015.* 78

Figure 6.30 *2D-scatter plot of the monthly mean cloud cover shown by CLARA A2 and SYNOP (top) and the histogram of the difference between CLARA A2 and SYNOP (bottom) for the entire period.* 79

Figure 6.31 *This figure shows the cloud cover comparison of CLARA-A2 afternoon satellites and MODIS collection 6 AQUA monthly means for the entire available time series 2002-2015. The top panel shows the difference plot, the panel below the time series and the bc-RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA-A2 left and MODIS right). The bottom left panel shows the 2D histogram of all data points in time and space and the bottom right panel the averaged zonal mean for CLARA-A2 in red and MODIS Aqua in blue.* 82

Figure 6.32 *Cloud top pressure comparison of CLARA-A2 afternoon satellites and MODIS collection 6 AQUA monthly means for the entire available time series 2002-2015. The top panel shows the difference plot, the panel below the time series and the bc-RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA-A2 left and MODIS right). The bottom left panel shows the 2D histogram of all datapoints in time and space and the bottom right panel the averaged zonal mean for CLARA-A2 in red and MODIS in blue.* 84

Figure 6.33 *Cloud cover comparison of CLARA-A2 and ISCCP monthly means for the entire available time series 1994-2009. The top panel shows the difference plot, the panel below the time series and the bc-RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA-A2 left and ISCCP right). The bottom left panel shows the 2D histogram of all datapoints in time and space and the bottom right panel the averaged zonal mean for CLARA-A2 in red and ISCCP in violet.* 86

Figure 6.34 *This figure shows the cloud top pressure comparison of CLARA A2 and ISCCP monthly means for the entire available time series 1994-2009. The top panel shows the difference plot, the panel below the time series and the bc-RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA A2 left and ISCCP right). The bottom left panel shows the 2D histogram of all datapoints in time and space and the bottom right panel the averaged zonal mean for CLARA A2 in red and ISCCP in violet.* 89

Figure 6.35 *Cloud cover comparison of CLARA-A2 afternoon prime satellites and PATMOS-x monthly means for the entire available time series 1982-2014. The top panel shows the difference plot, the panel below the time series and the bc- RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA-A2 left and PATMOS-x right). The bottom left panel shows the 2D histogram of all data points in time and space and the bottom right panel the averaged zonal mean for CLARA-A2 in red and PATMOS-x in green.* 93

Figure 6.36 *Cloud top pressure comparison of results for CLARA-A2 afternoon satellites and PATMOS-x monthly means for the entire available time series 1982-2014. The top panel shows the difference plot, the panel below the time series and the bias-corrected RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA-A2 left and PATMOS-x right). The bottom left* *panel shows the 2D histogram of all data points in time and space and the bottom right panel the averaged zonal mean for CLARA-A2 in red and PATMOS-x in green.* 95

Figure 6.37 *Spatial distribution of afternoon CPH (expressed as liquid cloud fraction) from CLARA-A2 (a), PATMOS-x (b), Aqua MODIS (c) and ISCCP (d), averaged over the period when all data records were available (01/2003-12/2007).* 99

Figure 6.38 *Zonal average CPH for morning (a) and afternoon (b) satellites, for CLARA-A2, PATMOS-x, MODIS and ISCCP, computed from corresponding averages from their common periods (given in Figure 6.37 caption). The shaded area around CLARA-A2 curves denotes the target accuracy. Optimal accuracy is equal to 0.01.* 100

Figure 6.39 *Time series of the morning CPH from CLARA-A2, MODIS and ISCCP, averaged over the globe (a), the tropics (b) and the areas excluding the tropics (c). The darker and lighter shaded areas around the CLARA-A2 curves denote the optimal and target accuracies, respectively.* 101

Figure 6.40 *As in Figure 6.39, for afternoon satellites.* 102

Figure 6.41 *As in Figure 6.40, for CPH_Day.* 103

Figure 6.42 *As in Figure 6.38, for CPH_Day.* 104

Figure 6.43 *As in Figure 6.39, for the morning CPH_Day.* 104

Figure 6.44 *As in Figure 6.43, for the afternoon satellites.* 105

Figure 6.45 *Time series of the twilight CPH_Day from CLARA-A2 and PATMOS-x, averaged over the globe (a), the tropics (b) and the areas excluding the tropics (c). The lighter shaded areas around the CLARA-A2 curves denote the target accuracy, while optimal accuracy is 0.01. The spatial coverage (in %) of the averaged data in the global case is shown in (d).* 106

Figure 6.46 *Spatial distribution of the all-sky LWP from CLARA-A2 (a, b), MODIS (c, d), ISCCP (e, f) and PATMOS-x (g), separately for morning (left column) and afternoon (right column) satellites, averaged over the period when all data records were available (01/2003-12/2007). The all-sky LWP from PATMOS-x morning satellites was not available.* 110

Figure 6.47 *Zonal average all-sky LWP for morning (a) and afternoon (b) satellites, for CLARA-A2, PATMOS-x, MODIS and ISCCP, computed from corresponding averages from their common period (01/2003-12/2007). The darker and lighter shaded areas around the CLARA-A2 curves denote the optimal and target accuracies, respectively.* 111

Figure 6.48 *Time series of the morning all-sky LWP from CLARA-A2, MODIS and ISCCP, averaged over the globe (a), the tropics (b) and the areas excluding the tropics (c). The darker and lighter shaded areas around the CLARA-A2 curves denote the optimal and target accuracies, respectively.* 112

Figure 6.49 *As in Figure 6.48 but for the afternoon satellites.* 113

Figure 6.50 *Time series of the afternoon globally averaged all-sky liquid COT from CLARA-A2, PATMOS-x, MODIS and ISCCP (a), and corresponding results for liquid REFF (b).* 114

Figure 6.51 *The locations of the South Atlantic (S-Atl), South Pacific (S-Pac) and North Pacific (N-Pac) validation areas.* 115

Figure 6.52 *Time series of the monthly mean all-sky LWP from CLARA-A2 and UWisc for the period 1988-2012, over the southern Atlantic (a), the southern Pacific (b) and the northern Pacific (c), separately for morning and afternoon satellites. Corresponding biases are also shown. The shaded areas denote the optimal, target and threshold accuracies for the bias (dark, middle and light, respectively).* 117

Figure 6.53 *Spatial distribution of the all-sky IWP from CLARA-A2 (a, b), MODIS (c, d) and ISCCP (e, f), separately for morning (left column) and afternoon (right column) satellites, averaged over the period when all data records were available (01/2003-12/2007).* 119

Figure 6.54 *Zonal average all-sky IWP for morning (a) and afternoon (b) satellites, for CLARA-A2, MODIS and ISCCP, computed from corresponding averages from their common period (01/2003-12/2007). The darker and lighter shaded areas around the CLARA-A2 curves denote the optimal and target accuracies, respectively.* 120

Figure 6.55 *Time series of the morning all-sky IWP from CLARA-A2, MODIS and ISCCP, averaged over the globe (a), the tropics (b) and the areas excluding the tropics (c). The darker and lighter shaded areas around the CLARA-A2 curves denote the optimal and target accuracies, respectively.* 121

Figure 6.56 *As in Figure 6.55 but for the afternoon satellites.* 122

Figure 6.57 *Time series of the afternoon globally averaged all-sky ice COT from CLARA-A2, MODIS and ISCCP (a), and corresponding results for ice REFF (b).* 123

Figure 6.58 *Global JCH relative frequency distributions [colors, %] of CTP [hPa] and COT for all months during 2003-2014. The top row (panels a-c) are CLARA-A2, the middle row (panels d-f) are MODIS Collection 6, and the bottom row (panels g-i) are for PATMOS-x. Left column contains the JCHs over sea and land surfaces (sea+land), middle column over sea-only surfaces (sea) and right column over land-only surfaces (land). Histogram frequencies are normalized to unity, such that each histogram sums to 100%.* 125

Figure 6.59 *Same as in Figure 6.58, but for the tropics defined as 30°S to 30°N.* 126

Figure 6.60 *Same as in Figure 6.54, but for the southern hemisphere mid-latitudes defined as 30°S to 60°S.* 127

Figure 6.61 *Same as in Figure 6.54, but for the northern hemisphere mid-latitudes defined as 30°N to 60°N.* 128

Figure 6.62 *An example of AO index time-series and selected enhanced positive and negative phases of the AO oscillation. All events that exceed (fall below) one standard deviation AO index, shown by thin horizontal line, are considered as enhanced positive (negative) events. Similar criteria were used while selecting events during ENSO and IOD.* 131

Figure 6.63 *The monthly distribution of enhanced positive and negative oscillation events and the monthly normalization factors used to compute climatological means.* 131

Figure 6.64 *The spatial distribution of total cloud fraction anomalies (in %) observed in three data sets during enhanced positive (strong El Nino) and negative (La Nina) oscillation events. The pattern correlations of CLARA anomalies with MODIS in the tropics (30N-30S) are 0.98 and 0.97 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.88 and 0.96.* 132

Figure 6.65 *Same as in Figure 6.64, but for the IOD events. The pattern correlations of CLARA anomalies with MODIS in the tropics (30N-30S) are 0.91 and 0.93 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.88 and 0.92.* 133

Figure 6.66 *Same as in Figure 6.64, but for the AO events. The pattern correlations of CLARA anomalies with MODIS in the Arctic (60N-90N) are 0.58 and 0.72 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.52 and 0.43.* 134

Figure 6.67 *The spatial distribution of LWP anomalies (in g/m2) observed in the three data sets during enhanced positive (strong El Nino) and negative (La Nina) oscillation events. The pattern correlations of CLARA anomalies with MODIS in the tropics (30N-30S) are 0.63 and 0.65 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.36 and 0.22.* 135

Figure 6.68 *Same as in Figure 6.67, but for IWP anomalies. The pattern correlations of CLARA anomalies with MODIS in the tropics (30N-30S) are 0.79 and 0.82 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.50 and 0.63.* 136

Figure 6.69 *The time series of the bias between the CLARA A2 and the SYNOP cloud fractional cover monthly mean. The red line is the linear fit.* 137

Figure 6.70 *The monthly mean bias in total cloud fraction (CLARA-A2 minus MODIS C6) from 2003 till the end of 2014 for different regions across the globe. The Polar Regions contain areas with latitudes higher than 60 ^{o}, mid-latitude regions are between 30 ^{o}-60 ^{o} and the tropics 30 ^{o}S-30 ^{o}N. The grey, green and pink envelopes show threshold, target and optimal stability requirements respectively. The stability rate (in % per decade) in CLARA-A2 is shown in the top-left corner of each subplot.* 138

Figure 6.71 *Same as in Figure 6.70, but for cloud top pressure. The grey, green and pink envelopes show threshold, target and optimal stability requirements respectively. The stability rate (in hPa per decade) in CLARA-A2 is shown in the top-left corner of each subplot.* 139

Figure 6.72 *The monthly mean bias in total cloud fraction (CLARA-A2 minus PATMOS-x) from 1982 till the end of 2014 for different regions across the globe. The Polar Regions contain areas with latitudes higher than 60 ^{o}, mid-latitude regions are between 30 ^{o}-60 ^{o} and the tropics 30 ^{o}S-30 ^{o}N. The grey, green and pink envelopes show threshold, target and optimal stability requirements respectively. The stability rate (in % per decade) in CLARA-A2 is shown in the top-left corner of each subplot.* 140

Figure 6.73 *Decadal trends of CPH (a) and CPH_Day (c) bias between CLARA-A2 and MODIS (in fraction decade^{-1}), for afternoon satellites, estimated from all possible combinations of time periods equal or larger than 10 years. Corresponding time series of annual average biases are also shown (b and d).* 141

Figure 6.74 *Decadal trends of the all-sky LWP bias between CLARA-A2 and MODIS (in g m^{-2} decade^{-1}), separately from morning (a) and afternoon (c) satellites, estimated from all possible combinations of time periods equal or larger than 10 years. Corresponding time series of annual average biases are also shown (b and d).* 142

Figure 6.75 *As in Figure 6.74 but for the all-sky IWP.* 144

LIST of TABLES

Table 1.1 *Summary of validation results compared to target accuracies for each cloud product. Notice that accuracies are given as Mean errors or Biases (both terms being equivalent) valid for both negative and positive deviations. Results from consistency checks (not totally independent) are marked in blue.* 16

Table 1.2 *Summary of validation results compared to target precisions for each cloud product. Consistency checks marked in blue.* 17

Table 1.3 *Summary of validation results compared to target decadal stabilities for each cloud product. Consistency checks marked in blue.* 18

Table 3.1 *Spectral channels of the Advanced Very High Resolution Radiometer (AVHRR). The three different versions of the instrument are described as well as the corresponding satellites. Notice that channel 3A was only used continuously on NOAA-17 and Metop-1. For the other satellites with AVHRR/3 it was used only for shorter periods.* 23

Table 3.2 *Channel 3A and 3B activity for the AVHRR/3 instruments during daytime. Notice that the given time periods show the availability in the CLARA-A2 data record and not the true lifetime of the individual sensor/satellite.* 23

Table 4.1 *Contingency table for the 2x2 problem.* nij *is the number of cases where CLARA reports event i and the reference reports event j. For example event 1 may be* ***clear*** *and event 2 may be* ***cloudy****.* 27

Table 4.2 *CM SAF cloud products and their respective target requirements (defined in AD 1) for the GAC data record of level-2, level-2b and level-3 products. Notice that the requirement on mean error or bias for accuracy is valid for both negative and positive deviations.* 28

Table 5.1 *Cloud type categories according to the CALIOP Vertical Feature Mask product* 35

Table 5.2 *Some basic characteristics of the PATMOS-x retrieval methods* 36

Table 6.1 *Overview of reference data records used for the evaluation of CLAAS-2 level-2 parameters.* 40

Table 6.2 *Overview of all CALIPSO-CALIOP validation results for the CFC parameter. Black line divides the results between original (unfiltered) results to the left and filtered results to the right (using cloud optical depth 0.15 as filtering threshold). In the latter part we also include results for the additional year 2015.* 44

Table 6.3 *Overview of all CALIPSO-CALIOP validation results for the CFC parameter separated into morning and afternoon satellites. Black line divides the results between original (unfiltered) results to the left and filtered results to the right (using cloud optical depth 0.15 as filtering threshold).* 45

Table 6.4 *Overview of all CALIPSO-CALIOP validation results for the CMA-prob CFC parameter separated into morning and afternoon satellites. Black line divides the results between original (unfiltered) results to the left and filtered results to the right (using cloud optical depth 0.2 as filtering threshold).* 51

Table 6.5 *Compliance matrix of CFC level-2 and level-2b product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against CALIPSO observations applying a cloud optical thickness filter of 0.15 for the official CFC product and 0.20 for the CMA-prob product (see text for motivation for using different filters).* 54

Table 6.6 *Compliance matrix of found global CTH level-2 and level-2b product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against CALIPSO observations applying a cloud optical thickness filter of 1.0.* 55

Table 6.7 *Mean bias error (MBE), root mean square error (RMSE) and correlation coefficient (r-value) between CLARA-A2 and PATMOSX-x L2b daily cloud fraction.* 60

Table 6.8 *Same as in Table 6.7, but for CTP evaluation from CLARA-A2 L2b against PATMOS-x L2b.* 63

Table 6.9 *Overview of reference data records used for the evaluation of CLARA-A2 level-3 parameters.* 77

Table 6.10 *Compliance matrix of found global CFC monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against SYNOP observations.* 80

Table 6.11 *Compliance matrix of found global CFC monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against MODIS results (consistency check).* 83

Table 6.12 *Compliance matrix of found global CTP monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against MODIS results (consistency check).* 85

Table 6.13 *Compliance matrix of found global CFC monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against ISCCP observations (consistency check).* 88

Table 6.14 *Compliance matrix of found global CTP monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against ISCCP results (consistency check).* 91

Table 6.15 *Compliance matrix of found global CFC monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against ISCCP observations (consistency check).* 92

Table 6.16 *Compliance matrix of found global CTP monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against PATMOS-x results (consistency check).* 94

Table 6.17 *Data records, their version and instruments that were used for the evaluation of the CPP products.* 97

Table 6.18 *Overall requirement compliance of the CLARA-A2 CPH product with respect to the Mean Error and the bias-corrected RMS (bc-RMS). Consistency checks marked in blue.* 108

Table 6.19 *As in 6.18, for the CPH_Day product.* 108

Table 6.20 *Overall requirement compliance of the CLARA-A2 all-sky LWP product with respect to the Mean Error and the bias-corrected RMS (bc-RMS). Consistency checks marked in blue. Units are in g m^{-2}.* 118

Table 6.21 *Overall requirement compliance of the CLARA-A2 all-sky IWP product with respect to the Mean Error and the bias-corrected RMS (bc-RMS). Consistency checks marked in blue. Units are in g m^{-2}.* 123

Table 6.22 *Overview of reference data records used for the evaluation of CLARA-A2 level-3 decadal stability.* 137

Table 6.23 *Overall decadal stability requirement compliance of the Cloud Phase and Cloud Phase Day products. Units are in fraction $decade^{-1}$* 142

Table 6.24 *Overall decadal stability requirement compliance of the all-sky LWP bias against MODIS. Units are in g m^{-2} decade^{-1}.* 143

Table 6.25 *Overall decadal stability requirement compliance of the CLARA-A2 all-sky IWP product. Units are in g m^{-2} decade^{-1}.* 145

Table 7.1 *Summary of validation results compared to target accuracies for each cloud product. Notice that accuracies are given as Mean errors or Biases (both terms being equivalent) valid for both negative and positive deviations. Results from consistency checks (not totally independent) are marked in blue.* 148

Table 7.2 *Summary of validation results compared to target precisions for each cloud product. Consistency checks marked in blue.* 149

Table 7.3 *Summary of validation results compared to target decadal stabilities for each cloud product. Consistency checks marked in blue.* 150

1 Executive Summary

This CM SAF report provides information on the validation of the CM SAF GAC Edition 2 data records (to be officially named CLARA-A2) derived from the Advanced Very High Resolution Radiometer (AVHRR) observations onboard the NOAA satellites. The covered time period ranges from 1982 (first satellite NOAA-7) to 2015 (last satellite Metop-B).

This report presents an evaluation of the following products:

Fractional Cloud Cover	CM-11011 (CFC)
Joint Cloud property histogram	CM-11021 (JCH)
Cloud Top level	CM-11031 (CTO)
Cloud Phase	CM-11041 (CPH)
Liquid Water Path	CM-11051 (LWP)
Ice Water Path	CM-11061 (IWP)

An extensive validation of cloud products from the CM SAF GAC Edition 2 data record has been performed. The reference data records were taken from completely independent and different observation sources (e.g. SYNOP, CALIPSO-CALIOP, SSM/I and AMSR-E) as well as from similar satellite-based data records from passive visible and infrared imagery (MODIS, ISCCP and PATMOS-x). Studies were made based on a mix of level-2 and level-3 products, also addressing some specific aspects affecting inter-comparisons (e.g., cloud detection capabilities for very thin clouds). However, it should be noticed is that a somewhat larger emphasis has been put on the evaluation of level-2 products since these are now also official products in GAC Edition 2. More in depth inter-comparisons were also made with the PATMOS-x data record because of the close relation (being also based on AVHRR GAC data and using the same basic AVHRR FCDR).

Tables 1.1-1.3 below give an overview of all results with respect to the target accuracies, target precisions and requirements on decadal stabilities. How these results were derived and what assumptions and definitions that were used are outlined in detail in the specific sub-sections of this report.

Results show the following, product by product:

- **Fractional Cloud Cover (CFC)**

- The CM SAF GAC CFC product fulfils the Target requirements for both accuracy and precision when compared with all references
- The only exception can be seen for the precision of level-2 products compared with CALIPSO-CALIOP. However, we claim that this is due to an existing mistake in the current requirements (i.e., RMS values should be higher for level-2 products than for level-3 products).
- The requirement on decadal stability is fulfilled.

- **Cloud Top level (CTO)**

- The CM SAF GAC CTO level-3 product fulfils the Target requirements for all references except against MODIS
- The CM SAF GAC CTO level-2 product fulfils Threshold requirements and is very close to fulfilling also Target requirements.
- The requirement on decadal stability is fulfilled.

- **Cloud Thermodynamic Phase (CPH)**

- The CM SAF GAC CPH product fulfils optimal accuracy requirement against most references except against ISCCP, while the CPH-Day product always fulfils the target requirement.
- In both products, optimal precision requirement is fulfilled against most references except ISCCP, where target requirement is achieved.
- The target and threshold requirements for decadal stability are fulfilled for CPH-Day and CPH, respectively.

- **Liquid Water Path (LWP)**

- The CM SAF GAC LWP product fulfils optimal accuracy and target precision requirements with respect to the UWisc data set. Note that – as a consequence of necessary selections of the data – the validation with UWisc was restricted to oceanic, stratocumulus-dominated areas.
- Optimal accuracy requirement is fulfilled with respect to MODIS and PATMOS-x data records and threshold requirement is achieved with respect to ISCCP. Target precision requirement is achieved with respect to all data sets.
- The target requirement for decadal stability is fulfilled with respect to MODIS.

- **Ice Water Path (IWP)**

- The CM SAF GAC IWP product fulfils optimal accuracy requirements when compared with MODIS and ISCCP.
- Using the same data sets, target precision requirement is achieved.
- The target requirement for decadal stability is fulfilled with respect to MODIS.

- **Joint Cloud property Histograms (JCH)**

- This product is excluded from specific requirement testing because of being composed by two already existing products (COT and CTP)
- Nevertheless, the product has been inter-compared with corresponding results from ISCCP, MODIS and PATMOS-x showing many similarities but also some CLARA-A2 specific features.
- It is believed that the access to this product representation would greatly enhance the usefulness of the CM SAF GAC products in some applications (e.g., in climate model evaluation it is a central product for COSP simulators).

Table 1.1 *Summary of validation results compared to target accuracies for each cloud product. Notice that accuracies are given as Mean errors or Biases (both terms being equivalent) valid for both negative and positive deviations. Results from consistency checks (not totally independent) are marked in blue.*

Product		Accuracy requirement (Mean error or Bias))	Achieved accuracies
Cloud Fractional Cover	(CFC)	5 % (absolute)	-3.2 % (CALIPSO level-2) -3.1 % (SYNOP level-3) -4.9 % (PATMOS-x level-2b) -4 % (PATMOS-x level-3) -5.9 % (MODIS) -3.4 % (ISCCP)
Cloud Top Height Cloud Top Pressure	(CTH) (CTP)	800 m 50 hPa	-840 m (CALIPSO level-2) -4.3 hPa (PATMOS-x level-2b) -18 hPa (PATMOS-x level-3) -88 hPa (MODIS) 13 hPa (ISCCP)
Cloud Phase	(CPH)	10 % (absolute)	1-2 % (PATMOS-x) 1-5 % (MODIS) 1-9 % (ISCCP)
Liquid Water Path	(LWP)	10 gm^{-2}	-3.5 to -0.5 gm^{-2} % (UWisc) 4.3 gm^{-2} (PATMOS-x) -2.3 to 3.3 gm^{-2} (MODIS) 10 to 17 gm^{-2} (ISCCP)
Ice Water Path	(IWP)	20 gm^{-2}	0.6 to 5.1 gm^{-2} (MODIS) 7.4 to 8.6 gm^{-2} (ISCCP)
Joint Cloud Histogram	(JCH)	n/a	n/a

Table 1.2 *Summary of validation results compared to target precisions for each cloud product. Consistency checks marked in blue.*

Product		Precision requirement (bc-RMS)	Achieved precisions
Cloud Fractional Cover	**(CFC)**	**20 % (absolute)**	**40 %(CALIPSO level-2)** **6.7 % (SYNOP level-3)** **1.6 % (PATMOS-x level-2b/3)** **10 % (PATMOS-x level-3)** **7.3 % (MODIS)** **8.8 % (ISCCP)**
Cloud Top Height **Cloud Top Pressure**	**(CTH)** **(CTP)**	**1700 m** **100 hPa**	**2380 m (CALIPSO)** **11 hPa (PATMOS-x level-2b/3)** **85 hPa (PATMOS-x level-3)** **58 hPa (MODIS)** **93 hPa (ISCCP)**
Cloud Phase	**(CPH)**	**20 % (absolute)**	**6-7 % (PATMOS-x)** **8-9 % (MODIS)** **13-16 % (ISCCP)**
Liquid Water Path	**(LWP)**	**20 gm^{-2}**	**11-20 gm^{-2} (UWisc)** **17 gm^{-2} (PATMOS-x)** **8-11 gm^{-2} (MODIS)** **14-19 gm^{-2} (ISCCP)**
Ice Water Path	**(IWP)**	**40 gm^{-2}**	**20-24 gm^{-2} (MODIS)** **25-31 gm^{-2} (ISCCP)**
Joint Cloud Histogram	**(JCH)**	**n/a**	**n/a**

Table 1.3 *Summary of validation results compared to target decadal stabilities for each cloud product. Consistency checks marked in blue.*

Product		Decadal stability requirement (change per decade)	Achieved stabilities
Cloud Fractional Cover	(CFC)	2 % (absolute)	-1.3 % (SYNOP) n/a (CALIPSO) 0.2 % (PATMOS-x) -1.1 % (MODIS)
Cloud Top Height Cloud Top Pressure	(CTH) (CTP)	200 m 20 hPa	n/a (CALIPSO) -4.0 hPa (MODIS)
Cloud Phase	(CPH)	2 % (absolute)	1.6-2.2 % (MODIS)
Liquid Water Path	(LWP)	3 gm^{-2}	1.0-2.3 gm^{-2} (MODIS)
Ice Water Path	(IWP)	6 gm^{-2}	2.7-3.7 gm^{-2} (MODIS)
Joint Cloud Histogram	(JCH)	n/a	n/a

There are already several satellite-based climate data records available providing similar information. However, in our opinion the added value of the CM SAF data record is:

- Cf. MODIS: much longer record (34 years vs 13 years)
- Cf. ISCCP: more homogeneous (no GEO used) and more spectral channels used
- Cf. PATMOS-x: good to have two similar data records produced with different algorithms to identify strengths /weaknesses of both approaches
- Cf. CALIPSO, SSM-I, UWisc: difference measurement principles, different variables measured, longer time frame
- Availability of additional surface radiation and surface albedo products produced from the same original data

Finally, it should be emphasised that the CLARA-A2 processing effort included not only significant algorithm improvements but also an unprecedented and rigorous (compared to CLARA-A1) quality control procedure of the original AVHRR GAC level-1b data record. In this respect the new data record appears to be much more stable and robust compared to CLARA-A1 and even compared to data records such as PATMOS-x. This is also a consequence of the in-depth nature of all validation efforts and the execution of the imposed feedback loop recommended at the previous DRI-5 review for CLARA-A1. This has led to some delays in the processing but it has enabled early discovery and correction of some crucial weaknesses of both technical and scientific nature.

Further guidance on how to use the products is given in the product user manual [RD 1]. Basic accuracy requirements are discussed [AD 2] and defined in the product requirements document [AD 1], and the algorithm theoretical basis documents describes the individual parameter algorithms [RD 2 – RD 6]. References are also given to the algorithm theoretical basis document for the probabilistic cloud mask (demonstration product – [RD 7]) and to the validation report for the CM SAF CLAAS-2 data record [RD 8].

2 The EUMETSAT SAF on Climate Monitoring

The importance of climate monitoring with satellites was recognized in 2000 by EUMETSAT Member States when they amended the EUMETSAT Convention to affirm that the EUMETSAT mandate is also to “contribute to the operational monitoring of the climate and the detection of global climatic changes". Following this, EUMETSAT established within its Satellite Application Facility (SAF) network a dedicated centre, the SAF on Climate Monitoring (CM SAF, http://www.cmsaf.eu).

The consortium of CM SAF currently comprises the Deutscher Wetterdienst (DWD) as host institute, and the partners from the Royal Meteorological Institute of Belgium (RMIB), the Finnish Meteorological Institute (FMI), the Royal Meteorological Institute of the Netherlands (KNMI), the Swedish Meteorological and Hydrological Institute (SMHI), the Meteorological Service of Switzerland (MeteoSwiss), and the Meteorological Service of the United Kingdom (UK MetOffice). Since the beginning in 1999, the EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF) has developed and will continue to develop capabilities for a sustained generation and provision of Climate Data Records (CDR’s) derived from operational meteorological satellites.

In particular, the generation of long-term data records is pursued. The ultimate aim is to make the resulting data records suitable for the analysis of climate variability and potentially the detection of climate trends. CM SAF works in close collaboration with the EUMETSAT Central Facility and liaises with other satellite operators to advance the availability, quality and usability of Fundamental Climate Data Records (FCDRs) as defined by the Global Climate Observing System (GCOS). As a major task the CM SAF utilizes FCDRs to produce records of Essential Climate Variables (ECVs) as defined by GCOS. Thematically, the focus of CM SAF is on ECVs associated with the global energy and water cycle.

Another essential task of CM SAF is to produce data records that can serve applications related to the Global Framework of Climate Services initiated by the WMO World Climate Conference-3 in 2009. CM SAF is supporting climate services at national meteorological and hydrological services (NMHSs) with long-term data records but also with data records produced close to real time that can be used to prepare monthly/annual updates of the state of the climate. Both types of products together allow for a consistent description of mean values, anomalies, variability and potential trends for the chosen ECVs. CM SAF ECV data records also serve the improvement of climate models both at global and regional scale.

As an essential partner in the related international frameworks, in particular WMO SCOPE CM (Sustained COordinated Processing of Environmental satellite data for Climate Monitoring), the CM SAF - together with the EUMETSAT Central Facility, assumes the role as main implementer of EUMETSAT’s commitments in support to global climate monitoring. This is achieved through:

- Application of highest standards and guidelines as lined out by GCOS for the satellite data processing,
- Processing of satellite data within a true international collaboration benefiting from developments at international level and pollinating the partnership with own ideas and standards,
- Intensive validation and improvement of the CM SAF climate data records,

- Taking a major role in data record assessments performed by research organisations such as WCRP (World Climate Research Program). This role provides the CM SAF with deep contacts to research organizations that form a substantial user group for the CM SAF CDRs,
- Maintaining and providing an operational and sustained infrastructure that can serve the community within the transition of mature CDR products from the research community into operational environments.

A catalogue of all available CM SAF products is accessible via the CM SAF webpage, http://www.cmsaf.eu/. Here, detailed information about product ordering, add-on tools, sample programs and documentation is provided.

3 Introduction to the AVHRR GAC data record

Measurements from the Advanced Very High Resolution Radiometer (AVHRR) radiometer onboard the polar orbiting NOAA satellites and the EUMETSAT METOP satellites have been performed since 1978. Figure 3.1 gives an overview over all satellite observations for satellites carrying the AVHRR instrument in the period 1980-2016. The instrument only measured in four spectral bands in the beginning (AVHRR/1) but from 1982 a fifth channel was added (AVHRR/2) and in 1998 even a sixth channel was made available (AVHRR/3), although only accessible if switched with the previous third channel at 3.7 micron. Table 3.1describes the AVHRR instrument, its various versions and the satellites carrying them. The retrieval of cloud physical properties (in particular particle effective radius and liquid/ice water path) is sensitive to the shortwave infrared channel being used. Table 3.2 summarizes when either of the channels 3a and 3b have been active on the AVHRR/3 instruments. The AVHRR instrument measures at a horizontal resolution close to 1 km at nadir but only data at a reduced resolution of approximately 4 km are permanently archived and available with global coverage since the beginning of measurements. This data record is denoted Global Area Coverage (GAC) AVHRR data.

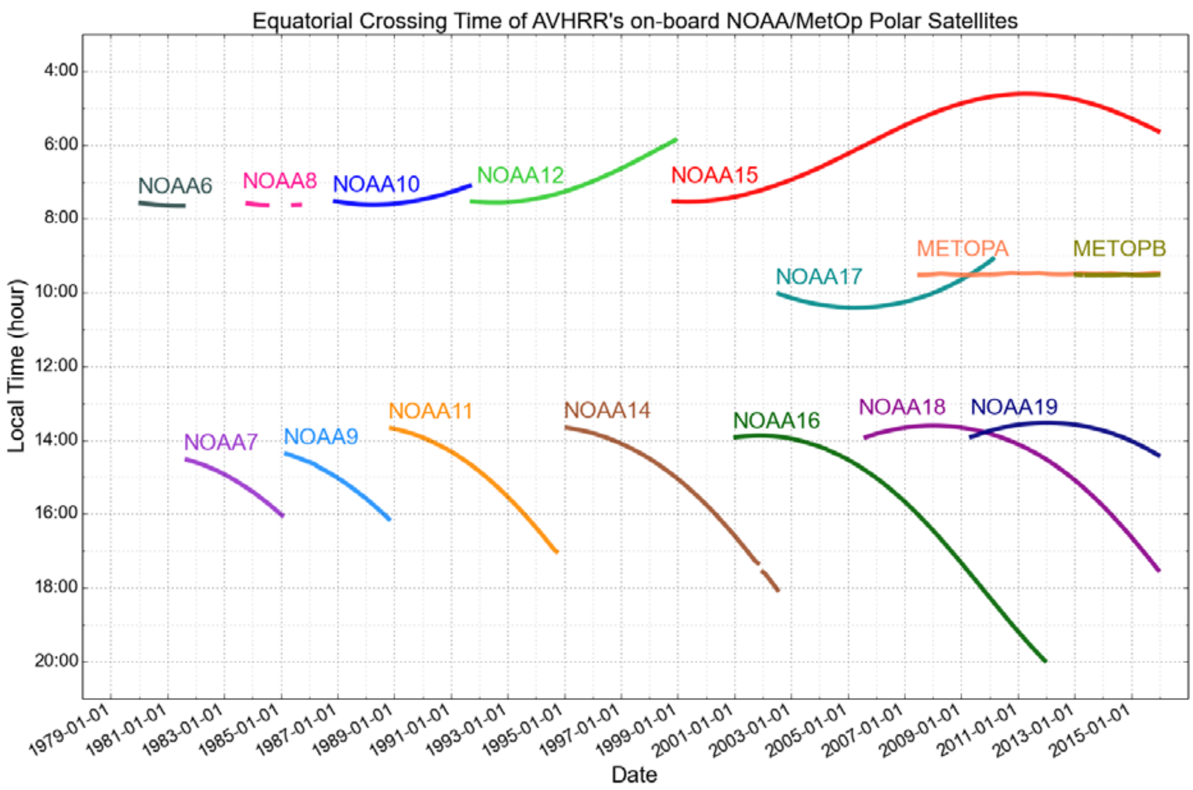


Figure 3.1 *Local solar times for daytime equator observations for all NOAA satellites from NOAA6 to Metop-B.*

Table 3.1 *Spectral channels of the Advanced Very High Resolution Radiometer (AVHRR). The three different versions of the instrument are described as well as the corresponding satellites. Notice that channel 3A was only used continuously on NOAA-17 and Metop-1. For the other satellites with AVHRR/3 it was used only for shorter periods.*

Channel Number	**Wavelength (micrometers) AVHRR/1** NOAA-6,8,10	**Wavelength (micrometers) AVHRR/2** NOAA-7,9,11,12,14	**Wavelength (micrometers) AVHRR/3** NOAA-15,16,17,18 NOAA-19, Metop-A, Metop-B
1	0.58-0.68	0.58-0.68	0.58-0.68
2	0.725-1.10	0.725-1.10	0.725-1.10
3A	-	-	1.58-1.64
3B	3.55-3.93	3.55-3.93	3.55-3.93
4	10.50-11.50	10.50-11.50	10.50-11.50
5	Channel 4 repeated	11.5-12.5	11.5-12.5

Table 3.2 *Channel 3A and 3B activity for the AVHRR/3 instruments during daytime. Notice that the given time periods show the availability in the CLARA-A2 data record and not the true lifetime of the individual sensor/satellite.*

Satellite	**Channel 3a active**	**Channel 3b active**
NOAA-15		06/1998 – 12/2015
NOAA-16	10/2000 – 04/2003	05/2003 – 12/2011
NOAA-17	07/2002 – 02/2010	
NOAA-18		09/2005 – 12/2015
NOAA-19		06/2009 – 12/2015
Metop-A	09/2007 – 12/2015	
Metop-B	01/2013 – 12/2015	

Figure 3.2 describes the actual coverage of observations in CLARA-A2 from each individual satellite over the entire period. Notice that the limitations to the use of AVHRR/2 and AVHRR/3 instruments leads to poorer time sampling (i.e., only one satellite available for daily observations) between 1982 and 1991. On the other hand, from 2001 and onwards more than two satellites are available for daily observations. The availability of observations peaks in 2009 where as many as six satellites are available (NOAA-15/16/17/18/19 + Metop-A). In the period 2010-2015 generally 5 satellites are available with the exception of 2012 with only 4 satellites.

To be kept in mind is that the CLARA-A2 data record was initially defined to cover the time period 1982-2014, i.e., the available AVHRR FCDR data record was prepared for that period. However, because of further delays in the processing it was decided to add also the year 2015 to the data record. Since results from 2015 are compiled using just an extrapolation of calibration curves for the visible AVHRR channels (i.e., no reference calibration measurements are available for this year as for the original FCDR data record), this calibration report will also evaluate results for 2015 separately for some core products (like CFC, the total cloud amount).

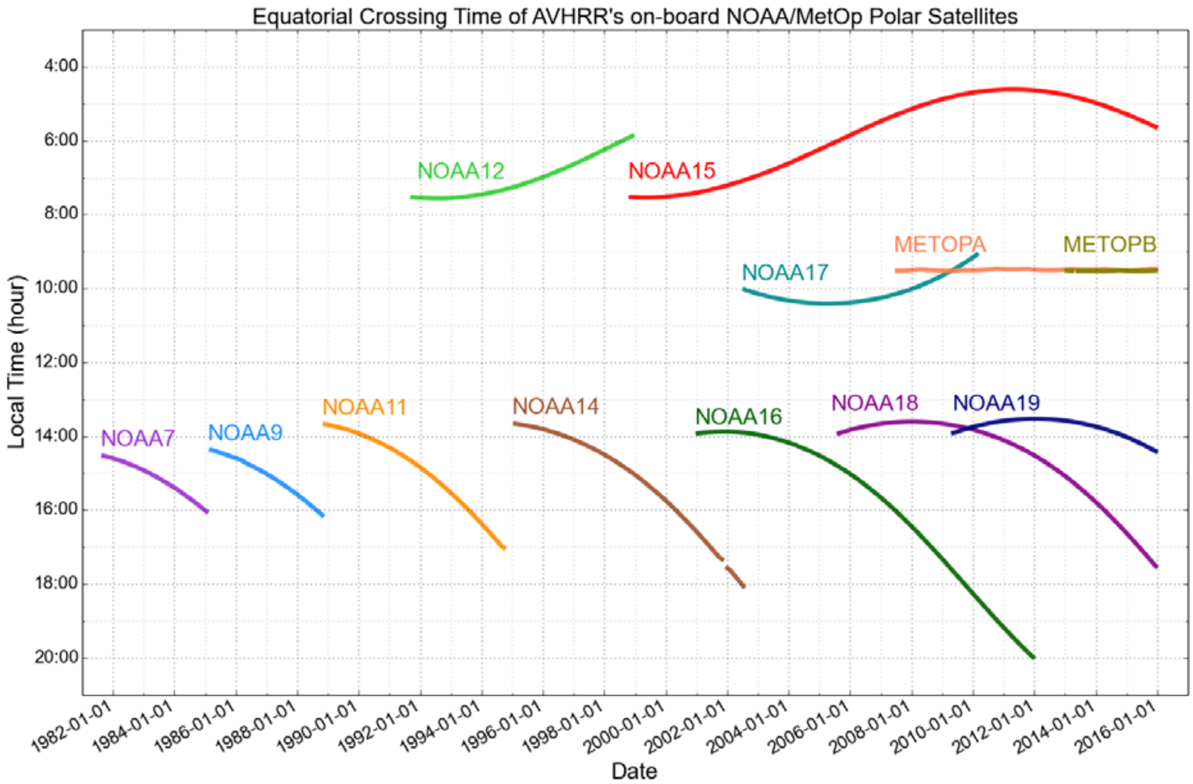


Figure 3.2 *Visualisation (same type as in Figure 3.1) of the used satellites in the CLARA-A2 data record.*

Observations from polar orbiting sun synchronous satellites are made at the same local solar time at each latitude band. Normally, satellites are classified into observation nodes according to the local solar time when crossing the equator during daytime (illuminated conditions). For the NOAA satellite observations, a system with one **morning observation node** and one **afternoon observation node** has been utilised as the fundamental polar orbiting observation system. This guarantees four equally distributed observations per day (if including the complementary observation times at night and in the evening when the satellite passes again 12 hours later). Equator crossing times have varied slightly between satellites. Morning satellites have generally been confined to the local solar time interval 07:00-08:00 and afternoon satellites to the interval 13:30-14:30. However, a change was introduced for the morning satellites NOAA-17, Metop-A and Metop-B, now being defined in a so-called mid-morning orbit with equator crossing times close to 10:00. A specific problem with the observation nodes for the NOAA satellites has been the difficulty to keep observation times stable for each individual satellite (Figure 3.1, described in more detail by Ignatov et al., 2004). No compensation for this has been attempted in the CLARA-A2 data record but corrections are considered for future CLARA versions.

This validation report describes the efforts to validate global cloud products retrieved by CM SAF cloud retrieval methods from AVHRR GAC data spanning the time period 1982-2015. Retrieval methods have been dependent on the access to two infrared (split-window) channels at 11 and 12 microns meaning that only data from satellites carrying the AVHRR/2 or AVHRR/3 instruments have been used.

An important aspect for any product-based climate data record (formally denoted Thematic Climate Data Records – TCDRs) is that retrieved products have been derived from accurately calibrated and homogenized radiances (formally denoted Fundamental Climate Data Records – FCDRs). For the CM SAF GAC data record we have used an AVHRR FCDR prepared by NOAA based on the work by Heidinger et al. (2010). This FCDR was prepared for the compilation of the "NOAA Pathfinder Atmospheres – Extended" (PATMOS-x) data record (for full description, see http://cimss.ssec.wisc.edu/patmosx/overview.html). The FCDR focusses in particular on the homogenization of the AVHRR visible reflectances and for CLARA-A2 we have used an updated calibration data record compared to the original one described by Heidinger et al. (2010). In addition to the prolongation of the covered period until 2015, the calibration method has been revised taking advantage of the new MODIS Collection 6 data record as its main calibration reference (publication currently in preparation). The calibration of infrared AVHRR channels is basically left untouched since the use of onboard blackbody calibration targets have been found to provide stable and reliable results. However, future upgrades of the AVHRR FCDR need to address remaining issues here also for the infrared channels (e.g., recognising the work of Mittaz et al., 2009).

4 Cloud products and validation strategy

In this report, we evaluate results for the following six cloud products derived from AVHRR GAC data (with formal product numbers and abbreviations according to AD 1 given to the right):

Fractional Cloud Cover	CM-11011	(CFC)
Joint Cloud property Histogram	CM-11021	(JCH)
Cloud Top level	CM-11031	(CTO)
Cloud Phase	CM-11041	(CPH)
Liquid Water Path	CM-11051	(LWP)
Ice Water Path	CM-11061	(IWP)

The theoretical basis for retrieval methods and compilation of TCDRs are described in RD 2. However, notice that RD 2 basically describes the methodology to prepare level-1, level-2/2b and level-3 data records while individual retrieval methodologies are described in RD 2-7.

The purpose of the validation effort is to evaluate whether products comply with product requirements stated in AD 1. These requirements are summarised further down in Table 4.2. The rationale for the chosen statistical parameters is that the overall SAF Product Requirements Table should include measures for both accuracy (i.e., how close to the truth is our estimation?) and precision (i.e., how stable is our estimation?).

For geophysical quantities at level-2, such as cloud top height, and for aggregated products (level-3), we use the bias, i.e. mean difference between CLARA and reference data as the metric for accuracy. In addition, the bias corrected root mean squared error (BC-RMSE) is used to express the precision of CLARA compared to a reference data record.

In case of discrete level-2 variables with only two possible events, e.g. cloud mask (*clear* or *cloudy*) and cloud phase (*liquid* or *ice*), we use the following scores which can be derived from the contingency Table 4.1.

- Probabilities of detection (POD) for event 1, 2: $\frac{n_{11}}{n_{11}+n_{21}}$, $\frac{n_{22}}{n_{22}+n_{12}}$

- False alarm ratios (FAR) for event 1, 2: $\frac{n_{12}}{n_{11}+n_{12}}$, $\frac{n_{21}}{n_{22}+n_{21}}$

- Hit rate: $\frac{n_{11}+n_{22}}{n_{11}+n_{12}+n_{21}+n_{22}}$

- Hanssen-Kuipers Skill Score (KSS): $\frac{n_{11}n_{22}-n_{21}n_{12}}{(n_{11}+n_{21})(n_{12}+n_{22})} \in [-1, 1]$

Table 4.1 *Contingency table for the 2x2 problem. n_{ij} is the number of cases where CLARA reports event i and the reference reports event j. For example event 1 may be **clear** and event 2 may be **cloudy**.*

	Reference reports 1	Reference reports 2
CLARA reports 1	n_{11}	n_{12}
CLARA reports 2	n_{21}	n_{22}

These scores can be viewed as measures of both accuracy and precision. However, they all have their specific advantages and disadvantages. Below we list some of the things that are typical and important to know about the various scores:

POD: The fraction of correct CLARA reports of a particular category relative to all ***reference reports*** of this category.

FAR: The fraction of incorrect CLARA reports of a particular category relative to all ***CLARA reports*** of this category.

Hit Rate: The total fraction of all correct CLARA reports (i.e., summing n_{11} and n_{22} in Table 4.1) relative to all ***reference reports***.

Kuipers Skill Score: This is a measure of correct CLARA reports, with random correct and unbiased reports subtracted out.

(*This is also a score that is better to use if one of the categories dominate. Thus, it punishes misclassifications of the minority category much harder than for other scores. However, a disadvantage is that it may be undefined in the case that there are only two of the four cases in Table 4.1 present. Perfectly correct CLARA reports are given the value 1 while totally opposite CLARA reports are given the value -1. Thus, results should preferably be higher than zero which represents totally random results.)*

Table 4.2 gives the target requirements for all CM SAF GAC cloud products. Observe that we describe two versions for the Cloud Top Level product (CM-11031) since we will use reference measurements made in pressure as well as in geometric altitude coordinates. In addition, there are no specific requirements given for the JCH product since it is composed by individual products COT and CTP. Table 4.2 only lists the target requirements for the accuracy and precision parameters. Compliance with a more relaxed threshold requirement and a more demanding optimal requirement (as defined in AD 1 and AD 2) are also discussed further in each specific sub-section for every cloud product. Regarding corresponding requirements for level-2 products, we notice that such requirements are generally similar to the requirements for the level-3 products. A useful guideline here is to consider that accuracy requirements should theoretically be very similar (at least if neglecting problems due to specific sampling methodologies) while precision requirements would generally differ (i.e., higher variability is expected for level-2 products).

Table 4.2 *CM SAF cloud products and their respective target requirements (defined in AD 1) for the GAC data record of level-2, level-2b and level-3 products. Notice that the requirement on mean error or bias for accuracy is valid for both negative and positive deviations.*

Product		**Accuracy requirement** (mean error = bias)	**Precision requirement** (bias-corrected RMS for CFC,CTH and CTP, RMS for all others)	**Stability requirement** (change per decade)
Cloud Fractional Cover	**(CFC)**	**5 % (absolute)**	**20 % (absolute)**	**2 % (absolute)**
Cloud Top Height	**(CTH)**	**800 m**	**1700 m**	**200 m**
Cloud Top Pressure	**(CTP)**	**50 hPa**	**100 hPa**	**20 hPa**
Cloud Phase	**(CPH)**	**10 % (absolute)**	**20 % (absolute)**	**2 % (absolute)**
Liquid Water Path	**(LWP)**	**10 gm^{-2}**	**20 gm^{-2}**	**3 gm^{-2}**
Ice Water Path	**(IWP)**	**20 gm^{-2}**	**40 gm^{-2}**	**6 gm^{-2}**
Joint Cloud Histogram	**(JCH)**	**n/a**	**n/a**	**n/a**

The requirement values listed in Table 4.2 are defined after taking into account requirements from different users and user groups. The most well-established reference here is the recommendations issued by the Global Climate Observation System - GCOS – community, see GCOS, 2006). However, values are also influenced by requirements from users working with regional climate monitoring and regional climate modelling applications (often having even stricter requirements than GCOS). More background on how the current requirements were established can be found in [AD 2].

The CM SAF GAC data record consists of instantaneous data (level-2 and level-2b), and daily and monthly mean (level-3) products for the period 1982-2015. Thus, the validation task is to evaluate the quality of these products. However, we also have to take into account that inter-comparing with level-3 products from other sources is much more difficult than to compare with instantaneous and simultaneous observations (i.e., the classical level-2 validation process). The reason is that level-3 products do not only depend on the quality of level-2 products but also on the method of compiling level-3 products (i.e., in terms of the applied temporal and spatial sampling, criteria for including or excluding a measurement, averaging method, etc.). This means that it is not always that level-3 product differences reflect true product differences in the same way as monitored by standard level-2 validation activities.

For practical reasons level-2 studies have been limited in time and space compared to the task of evaluating the full CLARA-A2 dataset. We believe that the mix of level-2 (instantaneous) and level-3 (monthly mean) studies provide enough information about the expected quality of daily level-3 products, which were not separately evaluated. The evaluation in this report is done with respect to ‘best practice’, based on the accessibility to high quality and homogeneous observations, which can be considered close to the truth and being independent, and based on well-established and highly utilized products.

A perfectly valid validation exercise requires access to high quality and homogeneous observations which can be considered close to the truth and being independent from the observations or measurements being evaluated. For a global data record of cloud products spanning a time period of 34 years these validation conditions do not exist, i.e., there is no high quality global observation data record that is covering the entire period in a homogeneous way. For that reason, we have been forced to use validation references that only partly fulfil the desired requirements.

The chosen validation references may be subdivided into two groups:

Group 1: Independent observations, which are generally considered to be true references, i.e. of superior quality. We have used the following observations:

- Cloud amount observations from surface stations (SYNOP)
  (time period 1982-2015)
- Cloud amount and cloud top observations from the CALIPSO cloud lidar (CALIOP)
  (time period 2006-2015)
- Cloud phase and ice water path from space-based lidar+radar DARDAR: January 2008
- Liquid water path from passive microwave sensors: 1988-2012

Group 2: Similar observation data records based on passive VIS-IR measurements, which are used for inter-comparisons rather than pure validation

- Cloud amount, cloud top, cloud phase, cloud phase and liquid/ice water path observations from the NOAA AVHRR Pathfinder Atmospheres – Extended (PATMOS-x) data record
  (time period 1982-2015)
- Cloud amount, cloud top, cloud phase and liquid/ice water path observations from the International Satellite Cloud Climatology Project (ISCCP)
  (time period 1982-2008)
- Cloud amount, cloud top, cloud phase and liquid/ice water path observations from the Moderate Resolution Imaging Spectroradiometer (MODIS)
  (time period 2000-2015)

Notice also that the evaluation of the joint histogram product (JCH) is based entirely on information provided by Group 2 above.

The first group of observations is definitely the most important group since it fulfils the condition that the observation reference must be independent. Thus, results achieved from comparisons with this group of observations will be given highest credibility.

However, as already stated, no reference is fulfilling the requirement of complete and homogeneous global and temporal coverage. Unfortunately, this concern especially group 1. It forces us to use other kind of reference data records to try to bridge existing gaps in the spatial and temporal domains, even if these data records cannot be considered as being completely independent. When dealing with the latter we also have to use (when available) existing knowledge of the quality of these data records. When such information is not easily found, we can at least try to utilize results from inter-comparisons with results from group 1 for the limited periods and spatial domains that are offered. In conclusion, results based on observations from reference group 2 should be considered as results from **consistency checks** rather than as results from a true validation effort. This will be pointed out repeatedly in the remainder of this report. We also conclude that for some products (CPH, LWP and IWP) we unfortunately must rely to a large extent on consistency checks since we do not have access to many completely independent observations.

The utilisation of the CALIPSO-CALIOP cloud observations in Group 1 above is worth a special statement. Despite the obvious limitations in both the temporal (i.e., only available for 10 years) and spatial (i.e., poor sampling since it only measures at nadir) domains, we are of the opinion that these observations must be utilised since they are probably the best cloud observations with global coverage that has ever become available. The idea has been to try to inter-compare with a limited but optimised CALIPSO data record to get the best possible information about the true CLARA-A2 performance of two of the products, namely CFC and CTH. This could then be put into relation with the results from all the other data records during the same limited period. Furthermore, these results should then be used as a baseline for the discussion of sub-sequent studies inter-comparing results for Group 2 for years before the CALIPSO observation period. For the future, we also believe that this optimised CALIPSO data record can serve as a tool for benchmark testing of new GAC Editions planned during the next CDOP phases.

The inter-comparison with PATMOS-x results has also a special position in this report, explaining the comparatively large share of the text. PATMOS-x is the only other data record using exactly the same fundamental input data (AVHRR GAC FCDR) as the CM SAF CLARA-A2 data record which makes comparisons natural. Regarding the analysis of the consistency checks for observations in Group 2 and the ability of making of a deeper analysis, we must state that only in the case of PATMOS-x we have had access to all underlying products (i.e., level-2b) so that more detailed analyses could be undertaken. In all other cases in Group 2 we only have level-3 data records (monthly means) which limits the further analysis to some extent.

In the following, we will first introduce in Section 5 the various reference data records we have used. Notice here that for each data record a special statement on errors and uncertainties is given at the end of the description. Section 6 presents validation results sub-divided into results for level-2 and level-3 products and also discussing in a third sub-section the decadal stability for all products.

5 Data Sets for Comparison with GAC

5.1 SYNOP: manual cloud observations from surface stations

Observations of total cloud cover made at meteorological surface stations (i.e. synoptic observations – hereafter called SYNOP) constitute one of the data records used to evaluate the cloud fractional coverage estimates. The SYNOP data used is from the local DWD archive of collected global SYNOP reports following the guidance of the *Guide to Meteorological Instruments and Methods of Observations* (WMO, 2008)

At manned stations the total cloud cover is visually estimated by human observers, at automated stations in contrast ceilometers are used for that purpose. For data quality reasons, only those SYNOP reports provided by manned airport stations were taken into account (~1800 stations globally).

SYNOP total cloud cover observations are used for the evaluation of level-3 cloud cover estimates.

Manual cloud observations are affected by many sources of error. We list some of the most important in the following:

- The observation is subjective in nature, i.e., despite clear instructions on how to make an observation, differences will appear because of different interpretations from person to person. This introduces a random noise in global cloud amount observations but may also lead to geographical biases (reflecting some systematic behaviour related to the way people have been educated/trained).
- The human eye has a detection limit for when a cloud can be clearly discernible against a cloud-free sky. This limit is somewhere in the cloud optical thickness range of 0.5-1.0 (with some dependence on solar zenith angle and on which viewing angles clouds are observed and the degree of aerosol load or haze in the troposphere). Thus, many satellite sensors have a higher sensitivity to e.g. cirrus detection than SYNOP observations.
- At night, the random error in the observations increases, naturally since the observer does not have a clear sky background against which a cloud can be observed (i.e., clouds are as dark as the cloud-free sky). However, accuracies improve in the presence of moonlight. Nevertheless, the overall effect is normally a negative bias (underestimated cloud amounts) since the observer is tempted to report cloud free conditions as soon as stars becomes visible, thus neglecting that large fractions of thin cirrus and other cloud types may still be present.
- A well-known deficiency of SYNOP observations is the scenery effect, i.e. overestimation of convective cloud towers at a slanted view (Karlsson, 2003). This effect is thus most pronounced in the summer season and for low to moderate cloud amounts when the overestimation easily can reach values of 20-30 % (1-2 octas).
- It is important to consider that most SYNOP stations are located at land stations and with higher density in developed countries. Thus, global averages tend to be biased towards land conditions in densely populated countries.

Since no rigorous study has been able to cover all those aspects in a quantitative manner (mainly because of lack of an absolute truth as reference) we can only make a very general qualitative statement about the overall quality. We would suggest that the accuracy of SYNOP observations vary between approximately +10 % (some overestimation) at daytime conditions changing to -10 % or worse (some underestimation) at night time. However, the variability (precision) probably reaches higher absolute values and it is largest during night conditions. This may lead to a strong seasonal variation with the worst accuracy and precision features during the winter season (at least at middle and high latitudes including the Polar Regions).

It is worth noting that the increasing trend to replace manual cloud observations with automatic observations from ceilometers will change the accuracy and precision of cloud observations in several ways. This may possibly lead to improved accuracies at night time but there is also a considerable risk that the precision figures degrades, mainly as an effect of that ceilometers only observer a very small fraction of the sky.

Despite their subjective character and varying quality, SYNOP observations still provide a useful reference data set suitable for monitoring and validating space-based estimations of cloud coverage, especially due to their long-term availability.

5.2 CALIPSO-CALIOP

Measurements from space-born active instruments (radar + lidar) provide probably the most accurate information we can get about cloud presence in the atmosphere. The reason is the fact that the measured reflected radiation comes almost exclusively from cloud and precipitation particles and is therefore not "contaminated" by radiation from other surfaces or atmospheric constituents as is the case for measurements from most passive radiometers. In this validation study we have decided to utilise measurements from the CALIOP lidar instrument carried by the CALIPSO satellite (included in the A-Train series of satellites - Figure 5.1).

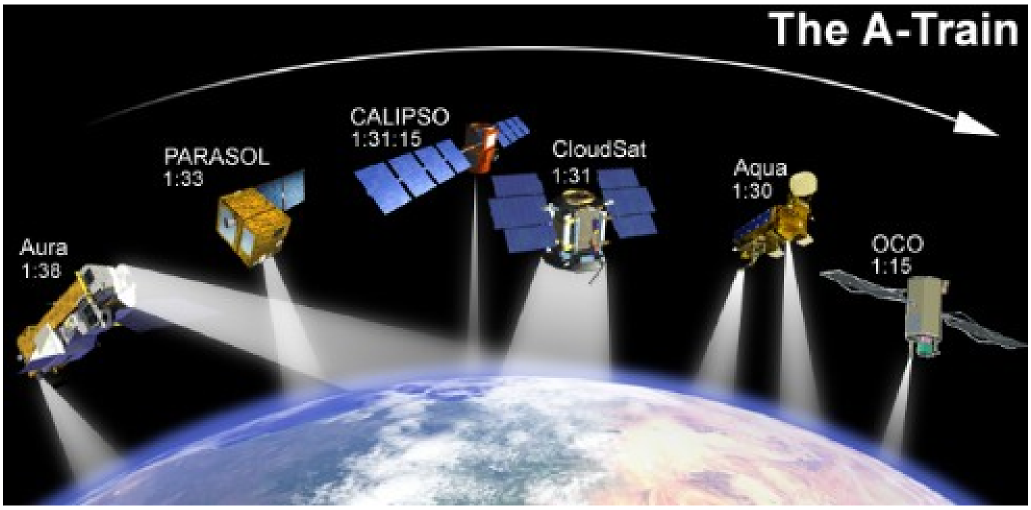


Figure 5.1 *The Aqua-Train satellites. (Image credit: NASA)*

The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite was launched in April 2006 together with CloudSat. The satellite carries the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) and the first data became available in August

2006 (Winker et al., 2009). CALIOP provides detailed profile information about cloud and aerosol particles and corresponding physical parameters (Vaughan et al., 2009).

CALIOP measures the backscatter intensity at 1064 nm while two other channels measure the orthogonally polarized components of the backscattered signal at 532 nm. The CALIOP cloud product we have used report observed cloud layers i.e., all layers observed until signal becomes too attenuated. In practice the instrument can only probe the full geometrical depth of a cloud if the total optical thickness is not larger than a certain threshold (somewhere in the range 3-5). For optically thicker clouds only the upper portion of the cloud will be sensed. The horizontal resolution of each single FOV is 333 m and the vertical resolution is 30-60 m.

The CALIOP products are available in five different versions with respect to the along-track resolution ranging from 333 m (individual footprint resolution), 1 km, 5 km, 20 km and 80 km. The four latter resolutions are consequently constructed from several original footprints/FOVs. This allows a higher confidence in the correct detection and identification of cloud and aerosol layers compared to when using the original high resolution profiles. For example, the identification of very thin Cirrus clouds is more reliable in the 5 km data record than in the 1 km data record since signal-to-noise levels can be raised by using a combined data record of several original profiles.

We used the CALIOP level-2 1 km and 5 km cloud layer data record versions 3-01, 3-02 and 3-30 (CALIPSO Science Team, 2015) for the validation purpose. The 5 km resolution data record is closest to the nominal AVHRR GAC resolution but according to Karlsson and Johansson (2013) there are some inconsistencies between results for the two resolutions which means that the total cloud amounts from the 5 km is often slightly underestimated. It means that some of the thick (opaque) boundary layer clouds that are reported in fine resolution (333 m and 1 km) data records are not reported in the higher resolution (5 km or higher) data records. This has to do with the methodology to do averaging at the longer scales (5 km or higher) where contributions from strongly reflecting clouds are removed from the original signal to facilitate detection of very thin cloud layers and aerosols. Thus, we use here the method proposed by Karlsson and Johansson (2013) combining the two CALIPSO data records (i.e., adding missed clouds at 5 km resolution which are detected at 1 km resolution). This normally gives almost 5 % higher global cloud amounts compared to if just relying on 5 km data (note: this estimation was made from data used in the validation study described in section 6.1.1.2) .

The CALIOP cloud layer product reports up to 10 cloud layers per column and provides information about cloud phase and cloud type of each layer as well as the pressure, height and temperature at each layer’s top.

The CALIOP data record classifies cloud layers into cloud types according to Table 5.1. To be noticed here is that the ISCCP cloud type method has been used in the sense that the vertical separation of Low (categories 0-3), Medium (categories 4-5) and High (categories 6-7) clouds is defined by use of vertical pressure levels of 680 hPa and 440 hPa. However, the separation of thin and thick clouds is made using the information on whether the surface or lower layers below the current layer can be seen by CALIOP.

Table 5.1 *Cloud type categories according to the CALIOP Vertical Feature Mask product*

Category 0	Low, overcast, thin (transparent St, StCu, and fog)
Category 1	Low, overcast, thick (opaque St, StCu, and fog)
Category 2	Transition stratocumulus
Category 3	Low, broken (trade Cu and shallow Cu)
Category 4	Altocumulus (transparent)
Category 5	Altostratus (opaque, As, Ns, Ac)
Category 6	Cirrus (transparent)
Category 7	Deep convective (opaque As, Cb, Ns)

We only give a quite general description of the CALIPSO data records in this section. The details concerning the actual use of the data records are elaborated further in the following sections 6.1.1.2 and 6.1.2.1.

It should be emphasized that the CALIOP measurement is probing the atmosphere very efficiently in the along-track direction since it is a nadir pointing instrument. Here, cloud dimensions down to the original FOV resolution (333 m) will be detected. However, it should be made clear that the across-track extension of the observation is still limited to 333 m. Thus, to compare CALIOP-derived results with the results of 4 km GAC AVHRR pixel data is not entirely consistent (i.e., CALIOP is only capable of covering the GAC pixel properly in one direction and not in the perpendicular direction). However, we believe that this deficiency is of marginal importance. Most cloud systems on the GAC scale will be detected, e.g., it is very unlikely to imagine elongated clouds with size and shapes below 0.3x4 km that might risk remaining undetected within a GAC pixel that coincides with a CALIOP measurement. Most clouds will have aspect ratios for the two horizontal directions that guarantee detection by CALIOP. However, it is also clear that in situations with scattered (sub-pixel) cloudiness within the GAC FOV some optically thick clouds may be detected by AVHRR cloud schemes while not being covered at all by CALIOP FOVs. Thus, some small bias between AVHRR and CALIOP observations due to this effect appears unavoidable.

It is important to consider that the CALIOP lidar instrument is much more sensitive to cloud particles than the measurement from a passively imaging instrument. It means that a significant fraction of all CALIOP-detected clouds will not be detected from imagers. This sensitivity difference also propagates into CPH and CTH, which will typically be sensed at a lower cloud layer by passive instruments compared to CALIOP (see e.g., Hamann et al., 2014). Thus, to get reasonable and justified results one should theoretically consider filtering out the contributions from the very thinnest clouds. We have applied this approach in this validation study, both in the study of cloud amounts (CFC) and cloud top heights (CTO).

The cloud detection efficiency with CALIOP is slightly different day and night because of the additional noise from reflected solar radiation at daytime that can contaminate lidar backscatter measurements. However, Chepfer et al. (2010) reports that this can introduce an artificial difference of not more than 1 % when comparing night time and daytime data.

In conclusion: Despite the fact that the CALIPSO cloud observations most likely are the best available cloud reference data record being released so far, we might still see a negative bias of a few percent in cloud cover when using exclusively the 5 km data record. However, in this validation effort we have tried to compensate for this effect by combining the 1 km and 5 km data records following Karlsson and Johansson (2013). Other errors, e.g. due to mis-interpretation of heavy aerosol loads as clouds, are in this respect of minor importance when judging the effect on full global orbits.

5.3 PATMOS-x

The most appropriate satellite-derived climatology to compare CLARA-A2 with is the PATMOS-x data record. The acronym stands for “AVHRR Pathfinder Atmospheres – Extended” and the corresponding cloud products have been derived using the CLAVR-x method (Clouds from AVHRR – Extended, see Heidinger et al, 2005, Pavolonis et al., 2005, Thomas et al., 2004 and Heidinger and Pavolonis, 2009). As for the CM SAF PPS method, AVHRR radiances in all available spectral channels have been used to derive global cloud and radiation products over the entire lifetime of the AVHRR sensor. Some basic information about the used methodology for the derivation of various parameters is given in Table 5.2. To notice is that the cloud screening methodology of CLAVR-x has undergone a substantial revision lately compared to the method described by the cited references. The previous multispectral threshold approach has been replaced by a probabilistic methodology (naïve Bayesian classifier – see Heidinger et al., 2012). We have compared CM SAF results against the results produced by this new method. This means we have compared to PATMOS-x version v05r03. The most up-to-date publication describing the PATMOS-x data record is provided by Heidinger et al. (2014).

Table 5.2 *Some basic characteristics of the PATMOS-x retrieval methods*

Product	Methodology
Cloud amount	Computed from results of a statistical naïve Bayesian cloud mask trained from CALIPSO-CALIOP cloud information
Cloud top level	Optimum Estimation (OE) retrieval
Cloud phase	Multi-channel test scheme
Cloud optical thickness	OE retrieval (with look-up tables as CM SAF but with different radiative transfer models and ice particle definitions)
Cloud effective radius	OE retrieval (with look-up tables as CM SAF but with different radiative transfer models and ice particle definitions)
Cloud liquid water path	Calculated from optical thickness and effective radius (Stephens’ parameterization – same as CM SAF)
Cloud ice water path	Calculated from optical thickness and effective radius (Stephens’ parameterisation – same as CM SAF)

The PATMOS-x data record is prepared exclusively as so-called level-2b products. This means that, for each satellite, data from all orbits during one day have been sub-sampled to produce only two global products per day valid for the nominal local solar time for both the descending (southbound) and ascending (northbound) observation nodes.

Due to the very close relationship between the CLARA-A2 data record and PATMOS-x, we will spend a substantial part of the validation report inter-comparing the results of the two data records.

5.4 ISCCP

The International Satellite Cloud Climatology Project (ISCCP) provides cloud properties over a period of more than 35 years (Rossow and Schiffer, 1991; Rossow et al., 1996; Rossow and Schiffer, 1999). This project was established in 1982 as part of WCRP to collect weather satellite radiance measurements (from geostationary and polar orbiting satellites) and to analyze them to infer the global distribution of clouds, their properties, and their diurnal, seasonal and inter-annual variations. The resulting data records and analysis products are being used to study the role of clouds in climate, both their effects on radiative energy exchanges and their role in the global water cycle. This project and its results are considered to be the state of the art today on what can be derived from routine weather satellite data. ISCCP is the only other existing TCDR for cloud physical property products (here we mean products CPH, LWP and IWP). However, it has the disadvantage that it is based on different satellite types – polar and geostationary – of which most of the latter do not contain the necessary narrow-band channels for accurate retrieval of LWP and IWP.

The production of ISCCP has recently been transferred to the National Centers for Environmental Information (NCEI) and a new high-resolution version of the data record (to be denoted ISCCP-H) is under production (see https://www.ncdc.noaa.gov/isccp). Unfortunately, the new data record is still not released which means that we have compared to the previous ISCCP-D2 version covering the period 1983-2008 as prepared in the Global Energy and Water cycle Experiment (GEWEX) database (Stubenrauch et al, 2012).

5.5 MODIS

MODIS (or Moderate Resolution Imaging Spectroradiometer) is an advanced imaging instrument onboard the Terra (EOS AM) and Aqua (EOS PM) polar satellites (see http://modis-atmos.gsfc.nasa.gov/index.html).

Both Terra and Aqua orbits around the Earth are sun synchronous. Terra passes from north to south across the equator in the morning (local solar time 10:30), while Aqua passes south to north over the equator in the afternoon (local solar time 13:30). Terra MODIS and Aqua MODIS are viewing the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands or groups of wavelengths.

Since the Terra and Aqua satellites passes in very similar orbits (at least the afternoon orbit of Aqua) as the NOAA and Metop-A satellites and since MODIS observes with as much as 36 spectral channels (including all the AVHRR-like channels), corresponding cloud products from MODIS should serve as a top quality reference for corresponding cloud products retrieved from AVHRR data. MODIS uncertainties are indeed expected to be somewhat smaller than what can be obtained with AVHRR retrievals. For example: multiple CO_2 channels allow a more accurate cloud-top height determination, additional shortwave channels allow better discrimination of (thin) cirrus and a more reliable retrieval of cloud optical properties over very bright surfaces. Otherwise, uncertainties should lie in the same ballpark as for CLARA-A2. The main limitation of MODIS is the relatively short duration of the observation period, starting in 2000. We have used the level-3 MODIS gridded atmosphere monthly global products - MOD08_M3 (Terra) and MYD08_M3 (Aqua). They contain monthly 1° × 1° degree grid average values of atmospheric parameters related to atmospheric aerosol particle properties, total ozone burden, atmospheric water vapor, cloud optical and physical properties, and atmospheric stability indices. Statistics are sorted into 1° × 1° degree cells on an equal-angle grid that spans a (calendar) monthly interval and then summarized over the globe. For this particular study we have used data from Terra & Aqua MODIS Collection 6.

5.6 DARDAR

To complement the picture drawn by the CALIOP lidar also CPR onboard CloudSat is considered. CPR is a nadir-looking cloud profiling radar in principle working like MIRA (see section 5.2) but sensing the atmosphere from above at 94 GHz. The instruments sensitivity is defined by a minimum detectable reflectivity factor of -30 dBZ and calibration accuracy of 1.5 dB. The minimum detectable reflectivity factor requirement was reduced to -26 dBZ when the mission was changed to put CloudSat into a higher orbit for formation flying in A-train.
The DARDAR data record (Delanoë and Hogan, 2008) provides the result from a synergistic retrieval method combining the measurements from the CALIOP lidar, the CLOUDSAT radar and the MODIS imager, all three elements of the A-Train satellite constellation. By combining these different measurements, consistent profiles of microphysical properties are retrieved based on the specific particle size (instrumental) wavelength sensitivities. The lidar signals for instance are sensitive to the particle surfaces in the line of sight (~radius2), which is dominated by the smaller particles in a particle size distribution (PSD) whereas the radar signals are sensitive to the square of the particle volume which is dominated by the larger particles in the PSD. When both signals are available the combined PSD sensitivities provide the best guess of extinction, effective particle radius and IWC. When only one of the signals is available, i.e. when the lidar is fully attenuated or when the particles are too small to be detected by radar, the DARDAR retrievals are based on the single instrument parameterizations. The optimal estimation framework used for this retrieval ensures a smooth transition from these different regimes. The DARDAR product has the vertical resolution of CALIOP (30/60 m) and a horizontal resolution given by the radar footprint (700m). This is in contrast to the comparison to the CALIOP data (Section 5.2) which has been averaged to 5 km wide layers before being compared to the CLAAS-2 data records. The DARDAR data for the current evaluation has been downloaded from the ICARE site: http://www.icare.univ-lille1.fr/projects/dardar/overview_dardar_cloud.
DARDAR data is used for level-2 evaluation of CPH and IWP (including ice COT and ice particle effective radius). Comparisons with this data record are affected by the same issues related to high ice clouds as discussed for CALIOP, i.e. DARDAR is much more sensitive to thin ice cloud than passive imagers.
In the lower part of the atmosphere, the reflectance of the surface affects the backscattered radar signal, so clouds may not be properly detected below 1 km distance to the surface.

5.7 UWisc: liquid water path observations from microwave imagers

Passive microwave imagers, such as the Special Sensor Microwave/Imager (SSM/I) series, can be used to retrieve column-integrated liquid water along with water vapour and surface wind speed. Because the microwave (MW) channels fully penetrate clouds, they provide a direct measurement of the total liquid (but not solid) cloud condensate amount. For precipitating clouds an estimate of the rain water path has to be made and subtracted from the total liquid water path to retrieve the cloud liquid water path.

For the CLARA-A2 LWP evaluation the University of Wisconsin (UWisc) MW-based LWP climatology (O'Dell et al., 2008) was chosen as an independent reference data record. The LWP climatology is based on retrievals from various microwave radiometer instruments, including the SSM/I series, the Tropical Rainfall Measurement Mission Microwave Imager (TMI), and the AMSR-E. The most recent version of the data record that was used for the evaluation (version 4) spans the years 1988 – 2012. Liquid water path estimates are reported to have an accuracy of 15% to 30% (O'Dell et al., 2008).

Two remarks have to be made regarding the validation. First, the MW LWP measurements are only possible over ocean, so the validation is restricted to marine clouds. Second, since the MW measurements are not sensitive to ice, care has to be taken to select for the validation only those CLARA-A2 grid cells with a sufficiently low monthly mean ice cloud fraction.

6 Evaluation of CLARA-A2 parameters

The presentation of the results has been subdivided into the following three sub-groups:

1. **Validation of AVHRR instantaneous (level-2 and level-2b) products**
2. **Validation of AVHRR level-3 products (including joint cloud histograms)**
3. **Evaluation of decadal product stabilities**

The situation is a bit special for level-2b products since no reference (except PATMOS-x) is defined in the same way, i.e., being globally sub-sampled for one local observation time. However, since level-2b observations have been sub-selected from original level-2 products having the best (i.e., lowest) satellite viewing angles we should expect validation results to be as good or actually better than corresponding validation results for level-2 products. In particular, results would be very similar to validation results based on CALIPSO-CALIOP for the afternoon NOAA-satellites (NOAA-18, NOAA-19) since they have been derived from near-nadir observation conditions. Thus, level-2b results will largely be assigned from the latter study and from direct comparisons with PATMOS-x results.

Each group is described in the following sub-sections 6.1-6.3.

6.1 Evaluation of AVHRR instantaneous (level-2 and level-2b) products

This section covers the evaluation of CLARA-A2 level-2 products and is organized according to Table 6.1.

Table 6.1 *Overview of reference data records used for the evaluation of CLAAS-2 level-2 parameters.*

Section	Reference observations	Parameters
6.1.1	Calipso	CFC, CTH
6.1.2	PATMOS-x	CFC, CTP, CTH, LWP, IWP
6.1.3	DARDAR (Cloudsat-Calipso)	CPH, IWP, (ice COT, ice REFF)

6.1.1 Evaluation against CALIPSO-CALIOP

Following the approach by Karlsson and Johansson (2013), we have conducted an extensive comparison with high-quality cloud observations from the CALIPSO-CALIOP sensor using data from the ten-year period 2006-2015.

We have adopted the following strategy for this study:

- Select the best complete matches (i.e., entire global orbits) between afternoon orbit satellites (i.e, NOAA-18 and NOAA-19) and A-Train/CALIPSO.

  Best means in this respect that the core selected AVHRR orbits would include a Simultaneous Nadir Observation (SNO, i.e., when satellite orbits cross) with a maximum observation time difference of 45 seconds. In addition, use also portions of adjacent orbits where 'simultaneous' observations along the track are available within a maximum time difference of 3 minutes.

- Select the best matches between morning orbit satellites (i.e, NOAA-17, METOP-A and METOP-B) and A-Train/CALIPSO.

  Best means here that the selected AVHRR orbits would have 'simultaneous' AVHRR-CALIOP observations available within a maximum time difference of 3 minutes. The coverage will be limited to two zones centered around latitudes 70° North and 70° South due to the almost perpendicular angle between the orbital planes of morning satellites and CALIPSO. Thus, collocations are much shorter and take place across the AVHRR swath instead of along the swath as for afternoon satellites.

- Compile statistics on both cloud amount (CFC) and cloud top height (CTH) for the total data record and for the two separate afternoon and morning orbit data records. In addition, compile also results (where applicable) for selected regions depending on latitude, surface conditions and illumination conditions (day, night, twilight).

- Since the very sensitive CALIOP lidar is observing some clouds that have to be considered as "sub-visible" to AVHRR observations, investigate and compile results in two modes:

  1. Results for all CALIOP-observed clouds.
  2. Results exclusively for those clouds which can be considered as detectable for AVHRR.

Consequently, following the selection criteria formulated here we carried out comparisons for all possible collocations from October 2006 to December 2015. This resulted in collocations for 10 000 satellite orbits distributed over 5820 afternoon orbits and 4180 morning orbits. 8553 orbits were collected for the period 2006-2014 while 1447 orbits were collected for the additional year 2015.

6.1.1.1 Defining detectable clouds from AVHRR

The very last point in the study strategy listed previously means that we must find a way of determining the group "AVHRR-detectable clouds" before we list the detailed results.
We propose a method for this by investigating the quantity Hit Rate (defined in section 4) in the following way:

1. Hit Rate can be calculated as a function of filtered cloud optical depths. The filtering means that all cases with clouds with a vertically integrated cloud optical depth less than a certain value will be treated as if there were no observed clouds.
2. For very small cloud optical depths this would give increasing Hit Rates since these thin clouds are normally not detected by AVHRR-based methods.
3. However, at some point (with increasing filtered optical depth) we will find an increasing number of cases where clouds were actually detected. So, if removing them (or rather, treating them as representing cloud-free CALIPSO observations) the Hit Rate will eventually start decreasing again.
4. We suggest defining the AVHRR-detectable cloud optical thickness limit as the filtered optical depth value where the Hit Rate finds its maximum.. This can conceptionally be described as the cloud optical depth where half of such cloudy cases would be detected. For higher values of filtered optical depth most clouds are detected while below this value of filtered optical depth a majority of clouds remains undetected.

Figure 6.1 shows the resulting relation between the Hit Rate and the filtered optical depth based on all results derived in the period 2006-2014.The figure also includes corresponding results for a limited validation data record for the previous CLARA-A1 data record.

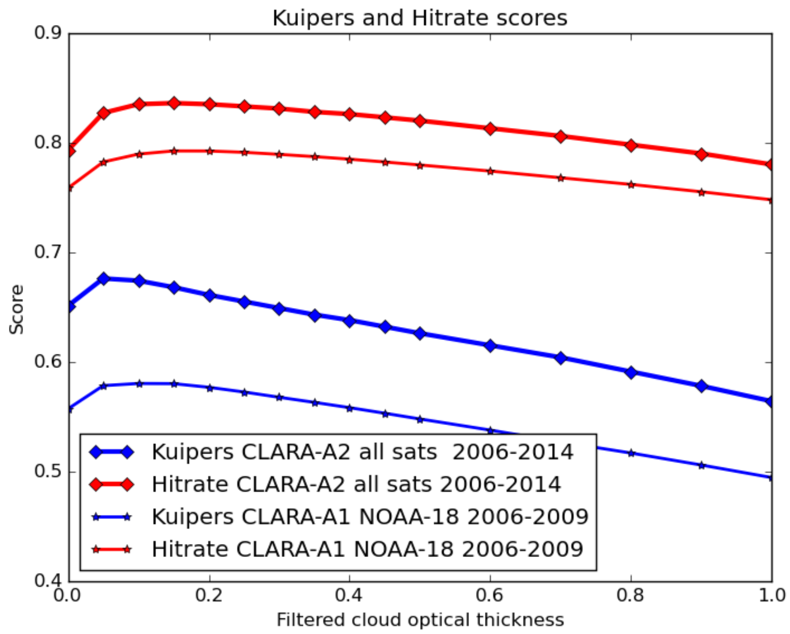


Figure 6.1 *Validation scores Kuipers and Hit Rate as a function of filtered CALIOP-estimated cloud optical depth (see text for explanation). Results are shown for CLARA-A2 in the period 2006-2014 and for CLARA-A1 for a limited validation data record with 99 NOAA-18 orbits 2006-2009.*

Closer examination of results in Figure 6.1 reveals that a peak Hit Rate value of 0.836 is reached for a filtered optical thickness of 0.15. We will use this as our lower cloud detection limit for when clouds are generally detected. Results derived at this cloud detection limit will also be referred to when assessing the fulfilment of product requirements.

As a final remark it can be noted from Figure 6.1 that, if choosing the Kuipers' score for finding the optimal performance, the lower cloud detection limit could be set to even lower cloud optical thickness values. This may be of importance if the desire is to clearly optimise the separability between cloudy and clear cases. However, for climatological purposes the Hit Rate seems more important since it has a more direct relation with overall mean cloud conditions.

6.1.1.2 Overall results for cloud fraction CFC

The following two tables summarise results using the validation scores previously defined in Section 4.

Results in Table 6.2 are given for the entire data record 2006-2014 compared to the limited validation data record for CLARA-A1. Here we also show results using the previously defined cloud detection limit at optical thickness 0.15 (section to the right of the black line). This also includes results for reference for the additional year 2015. We also show results sub-divided into morning and afternoon satellites in Table 6.3 This is motivated because of the different coverage of the Earth as explained earlier.

We first conclude from Table 6.2 that all validation scores have improved since CLARA-A1 except for the bias. However, the latter is explained by a considerably higher FAR(cloudy) for CLARA-A1 associated with a substantial portion of false clouds over semi-arid regions (explaining also the rather poor Kuipers score). In this context we should also mention that corresponding CLARA-A2 results (not shown here) for exactly the same set of 99 orbits used for CLARA-A1 validation do not differ significantly from the overall results seen here for the entire CLARA-A2 validation data record.

Filtered results in Table 6.2 show the optimal performance of CLARA-A2 after having reduced the influence of sub-visible clouds in the CALIOP data record. We notice that the bias has now been reduced to -3.2 % which is clearly within the target requirements of 5 % absolute. We also conclude that results for 2015 are only marginally different (e.g., a lower Kuipers score) which is mostly explained by the contribution to the results from the NOAA-18 satellite. This satellite has undergone a considerable orbital drift over the years which means that it is not any more perfectly aligned with the CALIPSO orbit. Thus, the average viewing angle at the collocations is not near-nadir any more but steadily increasing with time. This normally degrades results, since parallax effects leading to misprojected clouds become more and more important. However, if only looking at corresponding results for NOAA-19 (the other NOAA satellites in the afternoon orbit and still in a stable orbit) we find for 2015 a Kuipers score as good as 0.66 and a bias of -2.2 %. Thus, we conclude that results for 2015 are compatible with results for previous years in CLARA-A2. There is no sign of degraded results after using extrapolated calibration corrections for the visible channels.

Results in Table 6.3 for morning and afternoon satellites indicate a general superior performance of the afternoon satellites. However, this can largely be explained by a larger fraction of cases with polar winter conditions for the morning satellites because of the restriction to collocations near the latitude of 70 degrees. Thus, no collocations occur for morning satellites at low or medium latitudes where cloud detection conditions generally are more favourable. Also, collocations occur over a wide range of satellite viewing angles for morning satellites which means frequent misplacement of observed clouds due to parallax effects. We have not attempted any parallax corrections here since such corrections may introduce new collocation problems (e.g., requires assumption of homogeneous cloud layers extending over large areas).

Table 6.2 *Overview of all CALIPSO-CALIOP validation results for the CFC parameter. Black line divides the results between original (unfiltered) results to the left and filtered results to the right (using cloud optical depth 0.15 as filtering threshold). In the latter part we also include results for the additional year 2015.*

	CLARA-A1 unfiltered	CLARA-A2 unfiltered	CLARA-A1 filtered	CLARA-A2 filtered	CLARA-A2 filtered 2015
Number of orbits	99	8553	99	8 553	1447
Matched FOVs	725 900	24 345 199	725 900	24 345 199	2 440 092
Mean error (bias)	-14.4 %	-15.1 %	-4.2 %	-3.2 %	-3.0 %
RMS error (bias-corr)	44.7 %	40.2 %	45.2 %	40.2 %	42.3 %
POD cloudy	73.4 %	76.0 %	80.0 %	84.4 %	83.2 %
POD clear	82.3 %	89.1 %	78.0 %	82.5 %	79.9 %
FAR cloudy	8.4 %	4.7 %	14.3 %	11.1 %	12.6 %
FAR clear	45.9 %	44.1 %	29.7 %	24.1 %	26.1 %
Kuipers score	0.56	0.65	0.58	0.67	0.63
Hit Rate	75.8 %	79.3 %	79.2 %	83.6 %	82.0 %

Table 6.3 *Overview of all CALIPSO-CALIOP validation results for the CFC parameter separated into morning and afternoon satellites. Black line divides the results between original (unfiltered) results to the left and filtered results to the right (using cloud optical depth 0.15 as filtering threshold).*

	CLARA-A2 morning unfiltered	CLARA-A2 afternoon unfiltered	CLARA-A2 morning filtered	CLARA-A2 afternoon filtered
Number of orbits	3518	5035	3518	5035
Matched FOVs	1 648 967	22 696 232	1 648 967	22 696 232
Mean error (bias)	-19.5 %	-14.8 %	-8.1 %	-2.9 %
RMS error (bias-corr)	43.0 %	39.8 %	43.3 %	39.9 %
POD cloudy	69.0 %	76.5 %	77.3 %	84.9 %
POD clear	87.7 %	89.2 %	84.2 %	82.3 %
FAR cloudy	6.1 %	4.6 %	11.2 %	11.0 %
FAR clear	49.5 %	43.6 %	30.6 %	23.6 %
Kuipers score	0.57	0.66	0.61	0.67
Hit Rate	74.0 %	79.7 %	79.9 %	83.9 %

6.1.1.3 In depth analysis of results for cloud fraction CFC

The extensive character of the CALIPSO-CALIOP validation study offers in depth studies of e.g. regional, daily and seasonal performances of cloud detection. Such results will be reported (a manuscript for a peer-reviewed journal is under preparation) but we will here only highlight a few of these results accompanied with some selected illustrations.

We first notice that the high number of collocations in Table 6.2 (more than 24 million collocations for 10 000 orbits) over a time period of 10 years actually means that we have validation data with a good coverage of the entire Earth (although predominantly enabled from matchups with the afternoon satellites). It also means that we have a relatively good coverage of the global cloud conditions from CALIPSO-CALIOP during the last decade

(2006-2015). Thus, we can now try to rearrange and calculate the results in a global equal-area grid. We have used a Fibonacci grid with 28878 grid points evenly spread out around the Earth approximately 150 km apart. The resulting grid has almost equal area and almost equal shape of all grid cells. Fibonacci grids behave the same near the poles as at the equator, compared to traditional latitude-longitude grids which often behave in a strange way near the poles. For further details on Fibonacci grids, see González (2009) for pseudo code and comparison between Fibonacci and ordinary lat/lon grids and Swinbank and Purser (2006) for Numerical Weather Prediction applications.

Using the Fibonacci grid representation we now get a good approximation of the global mean CFC from CALIPSO-CALIOP for all the collocations with afternoon satellites (having full global coverage) for this period in Figure 6.2 Corresponding results from PPS cloud masks (i.e., the underlying cloud mask for CLARA-A2) are shown in Figure 6.3. We conclude that the agreement is striking and that only closer examination of certain areas (e.g. Polar Regions) shows some discernible differences.

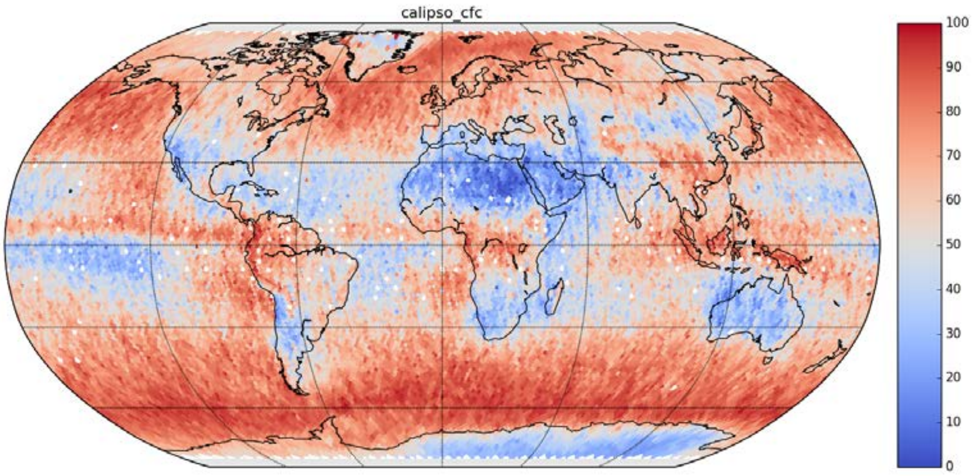


Figure 6.2 *Global mean CFC (%) from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations. White spots are positions with too limited coverage.*

The global gridding approach also means that for every listed validation score in the previous Table 6.2 and Table 6.3 we can now also plot the global distribution in a similar way which helps us understanding the global variability of the scores. To illustrate, Figure 6.4 shows the corresponding plot for the Hit Rate parameter while Figure 6.5 shows the Kuipers' score.

We can see that the two scores show two different sides of the performance, both important in their own way. For the Hit Rate we generally notice high values (e.g. 80 %) except over Greenland and Antarctica and the eastern parts of Africa and South America. We notice also that generally the score is high where cloud amounts are very high (e.g. mid-latitudes,

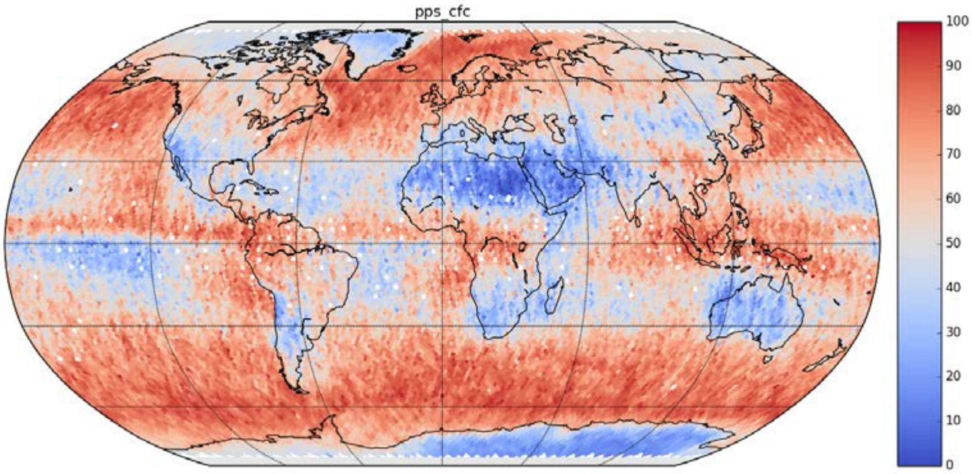


Figure 6.3 *Same as Figure 6.2 but calculated from PPS cloud masks (i.e., CLARA-A2).*

especially over ice-free ocean) or where cloud amounts are very low over land (e.g., over Sahara and Australia). This gives confidence in that the global extremes are properly covered. On the other hand, even if the Kuipers' score show almost a similar picture, we notice some particular problem areas like the Arabian Peninsula and Indonesia. For the Arabian Peninsula it is clear that the rare cases with clouds in this generally cloud free area are not optimally detected by CLARA-A2. In a similar manner it is clear that the detection of cloud-free portions in generally cloudy areas (e.g. mid-latitude oceans) are not optimally detected. However, in this case the deviation might come from CALIOP collocation problems (i.e. the occurrence of small-scale fractional clouds is high in these regions) rather than from real misclassifications. The collocation problem linked to small-scale clouds may also explain the somewhat reduced Hit Rate values in Figure 6.4 *Global distribution of the Hit Rate parameter from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations. White spots are positions with too limited coverage.* for the oceanic sub-tropical high locations

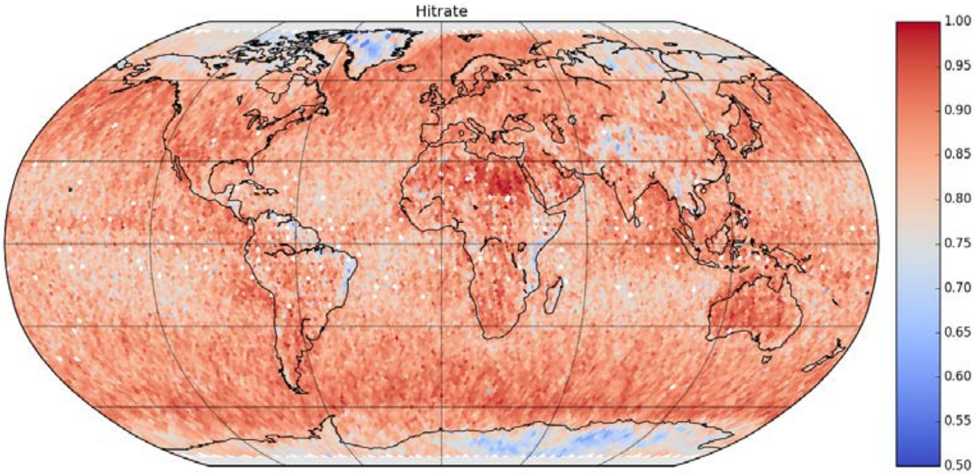


Figure 6.4 *Global distribution of the Hit Rate parameter from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations. White spots are positions with too limited coverage.*

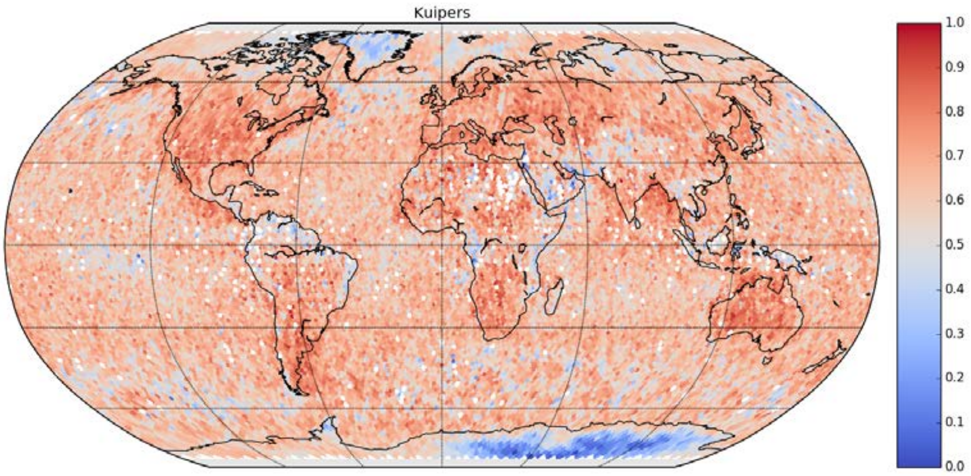


Figure 6.5 *Global distribution of the Kuipers score from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations. White spots are positions with too limited coverage to allow a confident definition of the Kuipers score.*

We may also use the global plots to illustrate which regions where our main target accuracy requirement (absolute bias of 5 %) is not fulfilled (Figure 6.6 *Global distribution of the CFC Bias from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations. Areas where Bias target requirements are not fulfilled are shown in red (excessive overestimation) and blue (excessive underestimation) colours.*. Here we clearly see that underestimated cloud amounts occur primarily over the Polar Regions but also over the southern part of the Eurasian continent. Excessive overestimation occurs primarily in the Tropical region (especially over Indian ocean and Indonesia) but also over some areas over high latitude oceans in the Northern Hemisphere (e.g., east of the US).

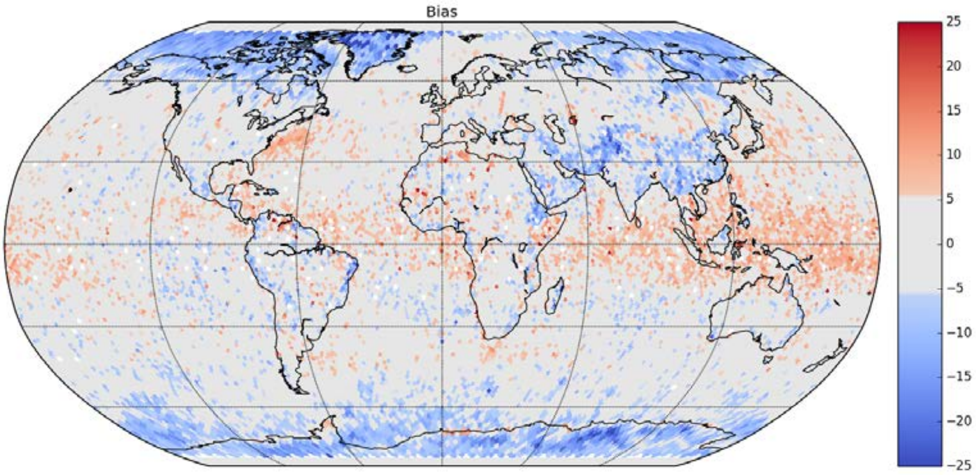


Figure 6.6 *Global distribution of the CFC Bias from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations. Areas where Bias target requirements are not fulfilled are shown in red (excessive overestimation) and blue (excessive underestimation) colours.*

The problems over the Polar Regions can be further illustrated by plotting results separately for the Polar areas during Polar night conditions (with complete darkness) in Figure 6.7 *Polar region distribution (Northern Hemisphere to the left and Southern Hemisphere to the right) of the probability of detecting clouds (PODcloudy) from CALIPSO-CALIOP calculated from **Polar night collocations** for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations.*. Here we choose the parameter POD(cloudy) since it explains most of the problems encountered here. We can clearly see that over all snow- and ice-covered parts of the Polar regions the POD(cloudy) values decrease considerably and in some places even reaching below the 50 % level (i.e., more than 50 % of all clouds remain undetected).

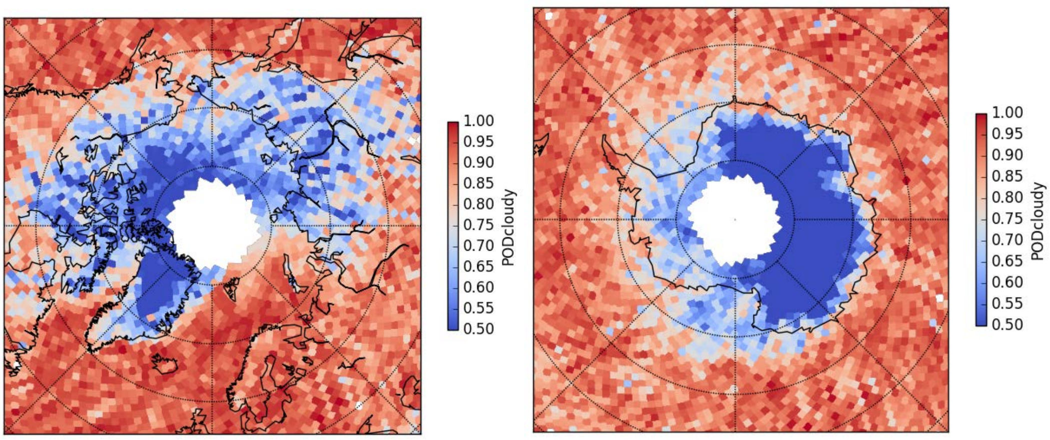


Figure 6.7 *Polar region distribution (Northern Hemisphere to the left and Southern Hemisphere to the right) of the probability of detecting clouds (PODcloudy) from CALIPSO-CALIOP calculated from **Polar night collocations** for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations.*

We conclude that AVHRR cloud detection in the Polar Regions still remains a big challenge. However, we would like to point out one strong and important feature of the CLARA-A2 data record: Cloud detection problems are limited to conditions at night (Polar Winter) and at twilight. During daytime (i.e., Polar Summer) cloud detection works much better and actually nearly as good as over any other place on Earth as illustrated in Figure 6.8. Only at very high-

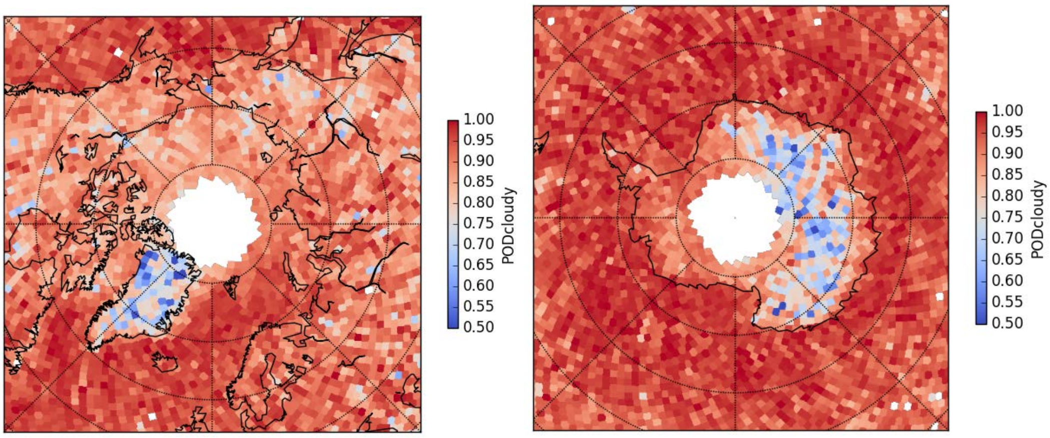


Figure 6.8 *Polar region distribution (Northern Hemisphere to the left and Southern Hemisphere to the right) of the daytime (Polar Summer) probability of detecting clouds (PODcloudy) from CALIPSO-CALIOP calculated from* ***Polar day collocations*** *for afternoon satellites in the period 2006-2014. Notice that results were derived using a cloud optical thickness limit of 0.15 for the CALIOP observations.*

altitude spots on Greenland and Antarctica we still see some problems. In that sense, the 34-year CLARA-A2 data record of high quality and well-validated cloud and surface radiation parameters (especially surface albedo) is probably unique.

Finally, a few more results should be highlighted:

- Global cloud detection performance is generally different depending on (solar) time of day. Cloud detection and cloud separation works best during daytime (Kuipers=0.70). Corresponding Kuipers values for night-time are 0.63 and for twilight 0.57.

- The occurrence of sub-visible cirrus clouds in the tropical zone (0-10° latitude) is higher than for other latitudes. This reduces unfiltered all-day Kuipers values to 0.64 compared to the value 0.69 for the sub-tropical zone (10-45° latitude).

- The current validation data record allows comparing the performance of daytime cloud detection between morning satellites and afternoon satellites over predominantly snow- and ice-covered surfaces in the Polar Regions (here defined as the area north/south of 75° latitude). This gives also a direct measure of how cloud detection works if using either the 3.7 micron channel (3b, afternoon satellites) or the 1.6 micron channel (3a, morning satellites). Results (from unfiltered data) are very clear here and clearly show that cloud detection works equally well for both constellations (i..e, Kuipers=0.64, Hit Rate=81 % for afternoon satellites and Kuipers=0.65, Hit Rate=81 % for morning satellites). Notice again that these scores during daytime in the Polar Regions are nearly as good as for any other area on Earth according to this study (see also Figure 6.8).

6.1.1.4 Validation results for CFC from probabilistic cloud masks

The CLARA-A2 level-2 product data record will contain a demonstration data record of probabilistic cloud masks (RD 7) for users who want to try using a more flexible cloud screening method. With such a cloud mask it is possible to use cloud mask confidence levels for applications that are very sensitive to any remaining misclassified clouds (e.g. SST retrievals). We have validated this product in the same way as the official CLARA-A2 level-2 cloud mask product. The goal has been to be able to as far as possible reproduce the results obtained by the standard cloud mask. Validation has been done as follows:

1. The probabilistic cloud mask (denoted CMA-prob) was originally trained against a limited CALIPSO data record of 99 collocated orbits (see Karlsson and Johansson, 2013). Here, the CALIPSO cloud mask was defined from the original cloud mask by thresholding the cloud optical thickness at the value 0.2. Consequently, we have here validated the cloud mask using the same threshold on the CALIPSO cloud product (i.e., filtering results using cloud optical thickness 0.2).
2. Even at a cloud optical thickness of 0.2 or slightly above, we cannot expect to detect all clouds from AVHRR. Thus, the derived cloud probabilities are likely to be slightly higher than the true probabilities because of some overtraining. This means that it is not meaningful at this stage to just accumulate results and inter-compare average cloud probabilities from CMA-prob and cloud frequencies from CALIPSO. Instead, it was shown by Karlsson and Johansson (2013) that the use of a cloud probability threshold of 60 % gave the best agreement with independent CALIPSO data records. Consequently, we will adopt the same method here and validate a resulting cloud mask created by thresholding the CMA-prob product at 60 % cloud probability.

Results in Table 6.4 show results sub-divided between morning and afternoon satellites (i.e., basically the same data record as in Table 6.3 but now for CMA-prob).

Table 6.4 *Overview of all CALIPSO-CALIOP validation results for the CMA-prob CFC parameter separated into morning and afternoon satellites. Black line divides the results between original (unfiltered) results to the left and filtered results to the right (using cloud optical depth 0.2 as filtering threshold).*

	CLARA-A2 CMA-prob morning unfiltered	CLARA-A2 CMA-prob afternoon unfiltered	CLARA-A2 CMA-prob morning filtered	CLARA-A2 CMA-prob afternoon filtered
Number of orbits	3520	5054	3520	5054
Matched FOVs	1 622 372	22 718 539	1 622 372	22 718 539

Mean error (bias)	-15.5 %	-17.3 %	-2.1 %	-3.7 %
RMS error (bias-corr)	46.4 %	40.6 %	46.4 %	40.6 %
POD cloudy	71.5 %	73.4 %	80.2 %	83.2 %
POD clear	79.7 %	90.0 %	75.8 %	83.2 %
FAR cloudy	9.4 %	4.5 %	16.8 %	11.4 %
FAR clear	49.5 %	46.4 %	28.1 %	24.0 %
Kuipers score	0.51	0.63	0.56	0.66
Hit Rate	73.7 %	77.6 %	78.4 %	83.2 %

By direct comparison with Table 6.3 we conclude that results are almost identical with only slightly better results for the official CLARA-A2 cloud mask. Results seem also more favourable for afternoon satellites than for morning satellites (e.g, FAR cloudy). The latter is probably related to the fact that the twilight category (which is more frequently occurring for the morning satellites having collocations only at high latitudes) is still not explicitly described by the CMA-prob methodology. The latter operates strictly in night or day mode while the NWCSAF PPS scheme applies a specific twilight thresholding sequence.

6.1.1.5 Validation of CTH validation results from CALIPSO-CALIOP

For the CTH validation we followed the same CALIOP-AVHRR matching procedure as for the CFC product and we used consequently the same collocated data record. Also here we applied cloud optical thickness filtering to the results but with a slightly different motivation and choice of thresholds. The main reason is that the radiance matching used to derive the cloud top height for passive sensors means that the derived cloud top would rather be representative for a height within the cloud layer itself ("the radiatively efficient height") than for the uppermost cloud top surface which is what the CALIOP measurement will report. Figure 6.9 illustrates how results change with changing value of the filtered cloud optical thickness. Here we also show results separately for cloud categories Low, Medium and High (following an ISCCP-type categorisation).

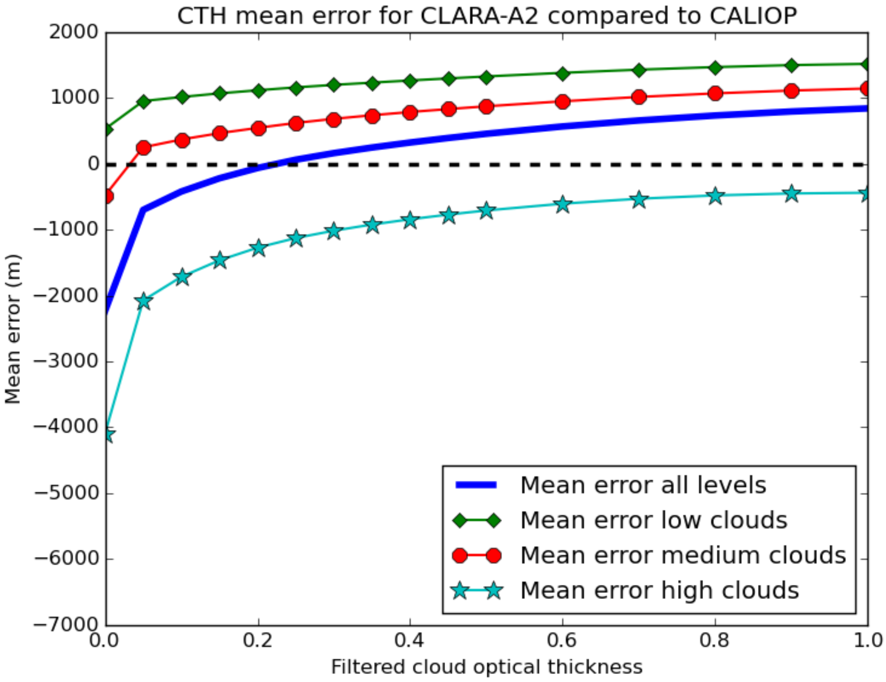


Figure 6.9 *Mean cloud top height (CTH) deviations from CALIPSO-CALIOP calculated from all collocations for afternoon satellites in the period 2006-2014. Results are given as a function of filtered cloud optical depths (see text for details).*

Regarding the meaning of the filtering it should be clarified that we remove clouds completely when we have an integrated cloud optical thickness which is less than the filtered threshold value. This is exactly identical to the case with evaluation of the CFC parameter. However, for the remaining clouds we also remove the uppermost layers from top-down until we reach the filtered cloud optical value. This means that we will systematically change upper-level clouds in multilayer cloud situations in the unfiltered case into Medium or Low clouds when we increase the filtered cloud optical thickness. For example, we notice in Figure 6.9 that increasing values of filtered cloud optical depth reduces the large underestimation of CTH for High clouds. This is expected and desired since we have a large contribution from semi-transparent Cirrus clouds where the effective cloud top height from the cloud radiance perspective should be much lower than the uppermost cloud top surface. However, at the same time it means that, if we have upper-level thin clouds superposed over Low or Medium clouds, the reclassification from High to lower cloud types will contribute to increasing the overestimation of these CTHs (because the high clouds contribute in making cloud top temperatures colder than the temperature of the true lower cloud layers). This leads to some confusion in the interpretation of the overall results. Nevertheless, if accepting the view that the measured radiance has contributions from several cloud layers the filtering procedure is theoretically reasonable. One could debate the most appropriate value to use for the filtering. The value should at least be larger than the corresponding value for evaluating cloud detection as in previous sections. An often used value in previous studies has been the value 1.0 (e.g., by Pincus et al., 2012) and we will use this value here to represent our final results in this report.

If first commenting the unfiltered results in Figure 6.9, it is clear that the CTH estimations of thin high level clouds becomes greatly underestimated if nothing is done to compensate for the influence of very thin clouds. However, if applying a filtering level of 1.0 the underestimation is reduced to a few hundred meters for High clouds. More alarming is that results for Low clouds in the unfiltered case yields overestimated cloud top heights of almost 600 meters. Thus, even in the case when there are no overlying high thin clouds present we have a considerable overestimation. This is a problem related to the very coarse vertical resolution and often weak strength of boundary layer inversions in the reference temperature profiles (ERA-Interim). Alternative methods to compensate for this have been developed (see Baum et al., 2012) but these have not yet been implemented in the PPS cloud retrieval scheme.

6.1.1.6 Summary of CFC and CTH validation results from CALIPSO-CALIOP

Table 6.5 summarises CALIPSO validation results for the official CLARA-A2 cloud mask and the probabilistic cloud mask CMA-prob. Corresponding results for the CLARA-A2 cloud top height product are given in Table 6.6.

For the CFC parameter in Table 6.5 the target requirements on the Bias parameter are fulfilled but not for the RMS parameter. However, as mentioned in section 4, the RMS should generally be higher for level-2 products and the current use of the same RMS requirement for level-2 and level-3 products is unfortunate and needs to be changed.

For the CTH parameter in Table 6.6 we have results that are close to fulfilling the target requirements on the Bias parameter. RMS values are not fulfilling target requirements but here we again argue (like for CFC) that the actual requirement values should be higher than for level-3 products.

Table 6.5 *Compliance matrix of CFC level-2 and level-2b product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against CALIPSO observations applying a cloud optical thickness filter of 0.15 for the official CFC product and 0.20 for the CMA-prob product (see text for motivation for using different filters).*

	CFC product requirements level-2 and level-2b			CLARA-A2 official CFC product (PPS cloud mask)	CFC based on Probabilistic cloud mask (CMA-prob)
	Threshold (absolute)	Target (absolute)	Optimal (absolute)		
Bias	10 %	5 %	2 %	-3.2 %	-3.0 %
bc-RMS	40 %	20 %	10 %	40 % (RMS)	43 % (RMS)

Table 6.6 *Compliance matrix of found global CTH level-2 and level-2b product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against CALIPSO observations applying a cloud optical thickness filter of 1.0.*

	CTH product requirements level-2 and level-2b			CLARA-A2 CTH
	Threshold	Target	Optimal	
Bias	1300 m	800 m	500 m	840 m
bc-RMS	3000 m	1700 m	1100 m	2380 m

6.1.2 Evaluation against PATMOS-x (level-2b)

The processing logic for the CLARA-A2 level-2b cloud products is described in RD 2. In this section, level-2b cloud products are evaluated against the latest (at the time of this report) PATMOS-x processed data record (denoted version v05r03). As described in RD 2, all level-3 cloud products are derived from the daily level-2b data record on a 0.05° equal-angle grid. Note that PATMOS-x level-2b cloud products are on an equal-angle grid of 0.1°. Any biases between CLARA-A2 and PATMOS-x level-3 cloud products (see section 5.3) should also emerge in the evaluation of level-2b products. For this reason, only the daily, area-weighted global evaluation of cloud fraction and cloud top pressure is performed for the overlapping time period of 1982 to 2014.

6.1.2.1 Inter-comparisons of daily CFC amounts in the period 1982-2014

Figure 6.10a shows the time series of the global, daily mean of cloud cover for CM SAF CLARA-A2 and for PATMOS-x computed from daily cloud fraction over the entire 1982-2014 observation period. Both data sets show generally stable cloud fractions, with a slight declining cloud fraction trend over the observation period; the decreasing trend is more apparent for PATMOS-x due to a systematic decrease in global cloud fraction occurring between years 1999 and 2003. The overall consistency in cloud fraction with time results in a relatively large correlation value of 0.72 (Figure 6.10a). The most evident feature of Figure 6.10a is the systematically lower cloud fractions for CLARA-A2 relative to PATMOS-x. The daily time series of the bias (CLARA-A2 minus PATMOS-x) is shown in Figure 6.10b. From 1982 to approximately 2003, CLARA-A2 cloud fractions ranged nearly 2 to 10% smaller than PATMOS-x. After 2003, a reduction in the mean bias is observed, reduced to a range of about -2 to -5%. This change in bias is likely associated with the change in number of AVHRR-carrying satellite observing platforms in orbit, which has increased considerably since the start of the 21st century (see Fig. 3.2); the increase in satellite observations also amounts to a broader temporal coverage over a single day, as there is a nearly equal presence of morning, afternoon, evening and early-morning/night observations. Over the full observation period, the mean bias error was -4.93% (Figure 6.10b) reducing to -3.87% for 2004-2014 and increasing to -5.45% for 1982-2003 (not shown). The reason for the reduction of the bias (which may appear strange if using exactly the same satellite data) in later years is related to the fact that the two methods have diurnal differences in cloud detection efficiency. This was earlier discussed for CLARA-A2 in Section 6.1.1.2. Results here indicate that PATMOS-x performs differently than CLARA-A2 over the course of the day. Thus, when the distribution of observations change over the day also the overall bias between the two data records may change.

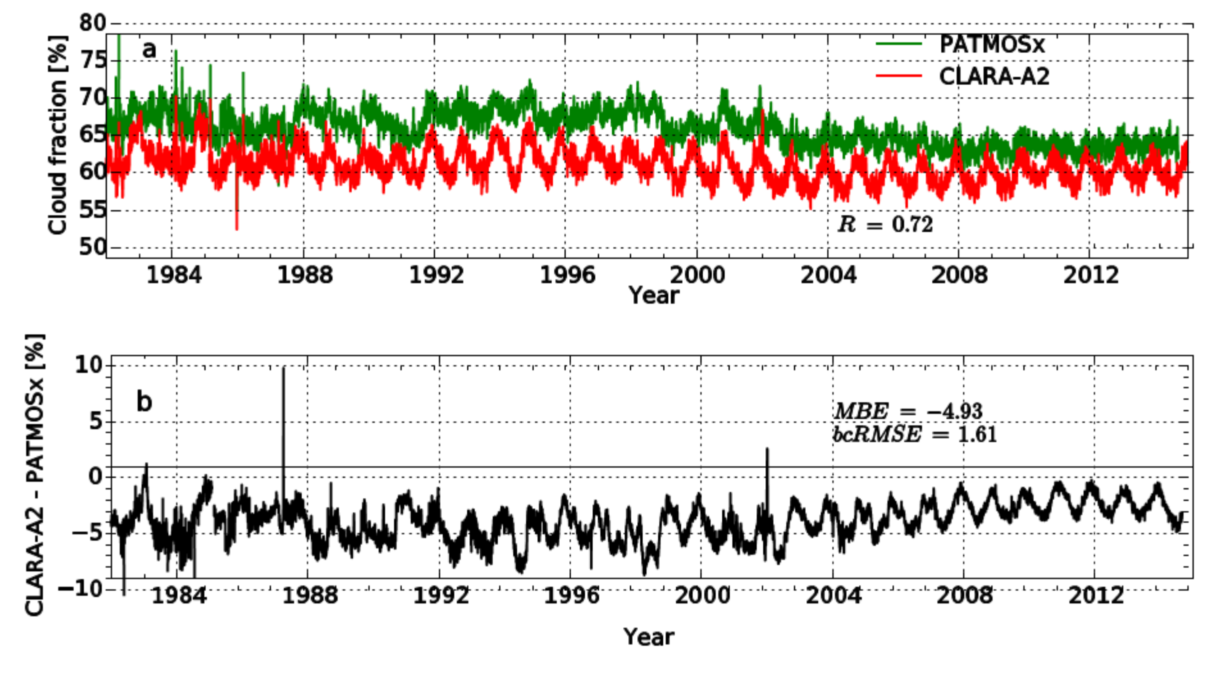


Figure 6.10 *a) Global mean cloud fraction [%] for PATMOS-x (green) and CLARA-A2 (red). Daily averages are computed from all (ascending + descending) satellite overpasses. The R-value (correlation coefficient) is provided in the lower right. Global averages are area-weighted. b) Daily global mean cloud fraction [%] difference, defined as CLARA-A2 – PATMOS-x. The mean bias error [%] and bias-corrected RMSE [%] in globally-averaged daily cloud fraction is provided in the upper right.*

It is apparent from Figure 6.10 that the biannual variation in yearly cloud fraction minima (late Northern Hemisphere summer) and maxima (late Northern Hemisphere winter) is considerably different between the two records. CLARA-A2 shows a much more pronounced biannual variation in globally-averaged seasonal cloud fraction compared to PATMOS-x. This biannual amplitude swing is generally consistent across the full record, whereas the PATMOS-x biannual cloud fraction maxima-minima tends to decrease in amplitude during the last decade.

Figure 6.11a shows 2D relatively frequency histograms of the cloud fraction for CLARA-A2 and PATMOS-x for the full data period with ascending and descending level-2b nodes combined. While some spread is apparent, the negative bias of CLARA-A2 cloud fraction relative to PATMOS-x is distinct; it is extremely rare for CLARA-A2 level-2b globally averaged cloud fractions to be larger than PATMOS-x, resulting in an RMSE of about 5.2%. The frequency distribution spread is reduced and a marked peak in cloud fraction underestimate of 2-5% is evident for ascending-only overpasses, which are primarily for sunlit scenes (Figure 6.11b). The descending nodes, which are primarily early evening or overnight scenes show a more frequent spread in globally daily cloud fraction, with a distinct distribution frequency closer to the 1:1 line (Figure 6.11c) compared with the ascending node distribution.

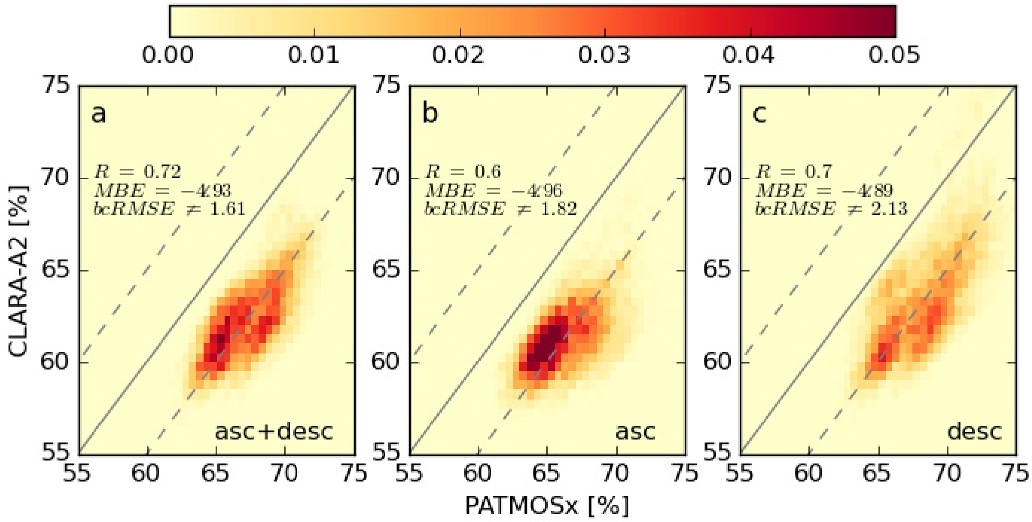


Figure 6.11 *2D relative frequency histograms for CLARA-A2 vs. PATMOS-x daily-averaged global cloud fraction from 1982-2014. Dashed lines show deviations from the 1:1 line of +5 % and -5 %.*

To examine the impact of overpass time further, daily global cloud fractions are shown separately only for the afternoon (local overpass approximately 13:30UTC – termed PM) and overnight (local overpass approximately 01:30UTC – termed AM) overpass nodes (Figure 6.12). Clearly the biannual variability in CLARA-A2 global, daily-averaged cloud fraction emerges during the overnight overpasses (Figure 6.12b-c); during the afternoon, CLARA-A2 daily cloud fraction appears much more stable, oscillating around 60% (Figure 6.12a). However, the sunlit scenes also appear to be affected by periods of satellite orbital drift or sensor stability, or both. This becomes apparent in the year ranges 1991-1994 and 1998-2002, where both PATMOS-x and CLARA-A2 show diverging cloud fraction trends (Figure 6.12a, c).

Despite the increased biannual cloud fraction variation for the AM overpasses, the correlation between PATMOS-x and CLARA-A2 is 15% higher than for the PM overpasses (Figure 6.12a-b); this is consistent with the increased correlation for all ascending and descending orbits (Figure 6.11b-c). Due to the larger annual variation in global cloud fraction, the mean bias error is approximately 0.5% larger (Figure 6.12a-b) (greater underestimation of CLARA-A2 relative to PATMOS-x). These results suggest that although the magnitude of seasonal cloud fraction variation differs, CLARA-A2 is in more agreement with PATMOS-x daily cloud fraction annual cycle for overnight observations when the global cloud fraction is largest (Figure 6.12a-b).

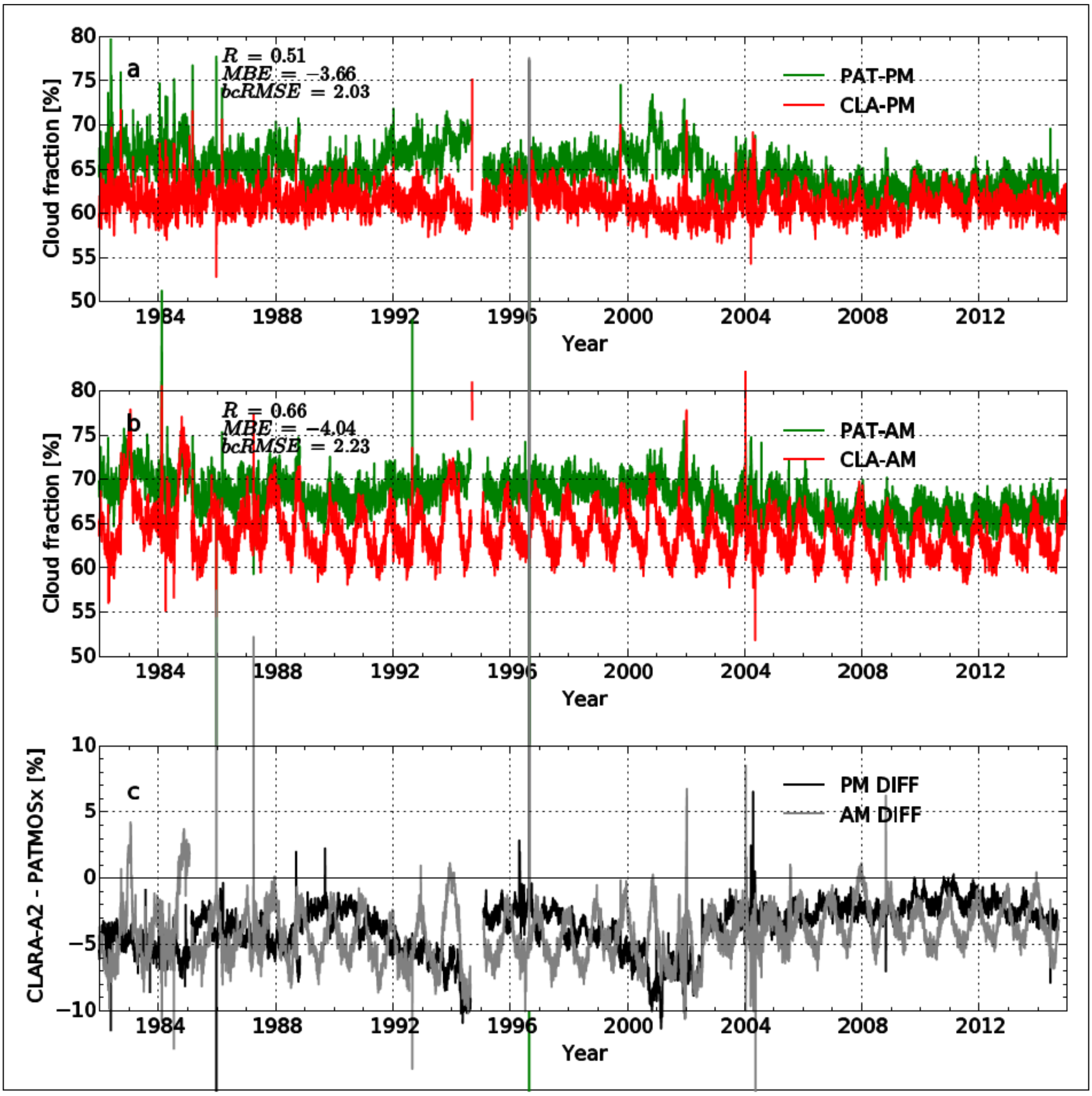


Figure 6.12 *Same as in Figure 6.10, but with global, daily-averaged cloud fraction for a) afternoon (PM) and b) overnight (AM) local satellite overpass times. c) Daily difference in global average cloud fraction for PM (black) and AM (gray) overpasses.*

A summary of the statistics of daily cloud fraction are provided in Table 6.7. A better correlation between the data records is observed for overnight observations, at the cost of a slight increase in CLARA-A2 cloud fraction underestimation; these features are related to the magnitude differences in biannual cloud fraction described above. Furthermore, excluding high latitudes pole ward of 60° results in a reduction of both MBE and RMSE of nearly 0.5 to 1% compared with the globally-averaged overpass nodes. This illustrates the complexity and difficulties in consistently masking cloudy pixels between the two data records over the Polar Regions.

Table 6.7 *Mean bias error (MBE), root mean square error (RMSE) and correlation coefficient (r-value) between CLARA-A2 and PATMOSX-x level-2b daily cloud fraction.*

Region / overpass	MBE (%)	RMSE (%)	r-value
Global / all	-4.93	1.61	0.72
Global / afternoon	-3.66	2.03	0.51
Global / overnight	-4.04	2.23	0.66
60°S-60°N / afternoon	-2.91	1.94	0.58
60°S-60°N / overnight	-2.95	1.73	0.81

Summary of results:

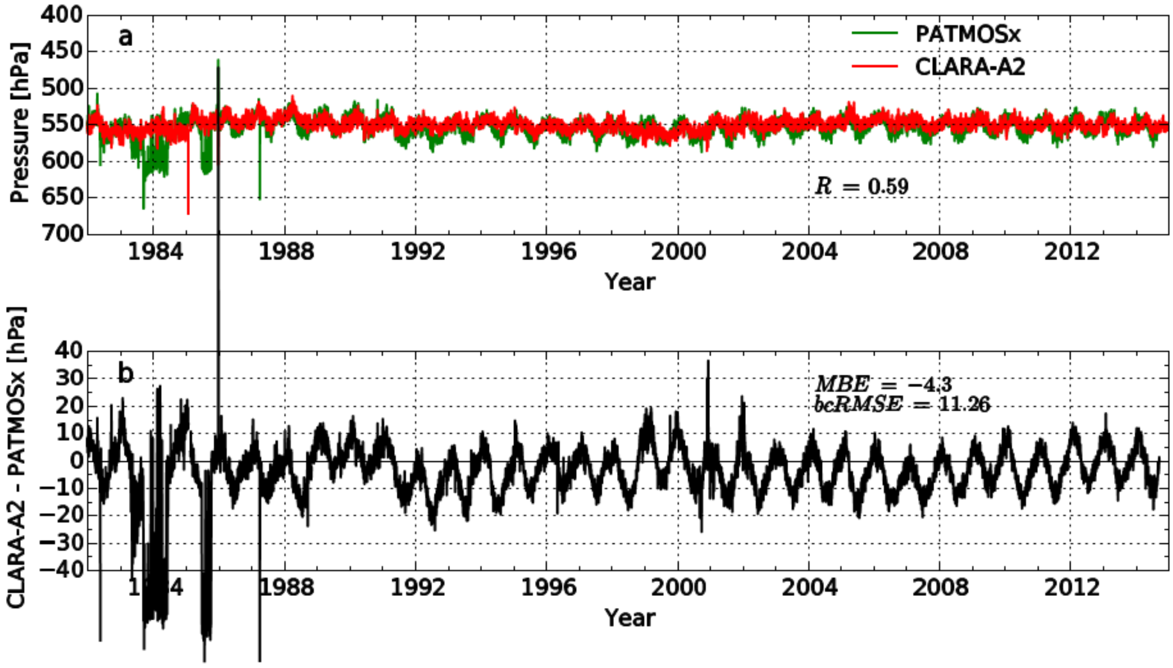
- Good agreement in global, daily-averaged cloud fraction resulting in a mean bias of -4.9% and a correlation of 0.72
- Cloud fraction time series for both CLARA-A2 and PATMOS-x level-2b are stable, indicating a weak globally-averaged declining cloud fraction trend from 1982-2014
- CLARA-A2 daily-averaged cloud fraction is systematically lower than PATMOS-x
- Biannual variation in global cloud fraction is largest for CLARA-A2. This results in a relative maximum in negative cloud fraction bias during late Northern Hemisphere summer, and a relative minimum in negative cloud fraction bias during late Northern Hemisphere winter
- The biannual cloud fraction variation is largest for overnight satellite observations, the biannual variation magnitude is relatively similar to PATMOS-x for afternoon satellite observations
- However, the correlation coefficients are larger for the overnight overpasses compared to the afternoon overpasses. This suggests that during the afternoon, there is a slight phase shift in the biannual cloud fraction variation

6.1.2.2 Inter-comparisons of daily CTP in the period 1982-2014

Level-2b cloud top pressure (CTP) is evaluated against PATMOS-x. We examine only the daily-average CTP, as well as the full-resolution CTP relative frequency distributions for January and July, over the years 1982-2014. A summary of the evaluation statistics for different regions and overpass nodes is provided.

Figure 6.13a shows the time series of global, daily-averaged CTP for PATMOS-x and CLARA-A2 while Figure 6.13b shows the same results but only for the common cloud mask (i.e., when both data records report cloudy conditions). Averaged global CTP is generally stable for both data records, with evidence of bi-annual variability in CTP. This results in a correlation of nearly 0.6 between the data records which increases to 0.8 for the common cloud mask case (thus, disagreeing cloud masks have some influence). Overall, there is a slight negative bias of 4.3 hPa in CLARA-A2 relative to PATMOS-x, meaning CLARA-A2 daily averaged cloud top height is slightly higher than PATMOS-x. This bias increases to 6.9

hPa for the common cloud mask case. After approximately 1991, the daily bias in CTP becomes generally consistent with the largest CTP underestimates (cloud top height overestimate) found during the Northern Hemisphere summer, peaking between -10 and -20 hPa (Figure 6.13c); the largest CLARA-A2 CTP overestimates occur during Northern Hemisphere winter and generally peak between +5 and +10 hPa.



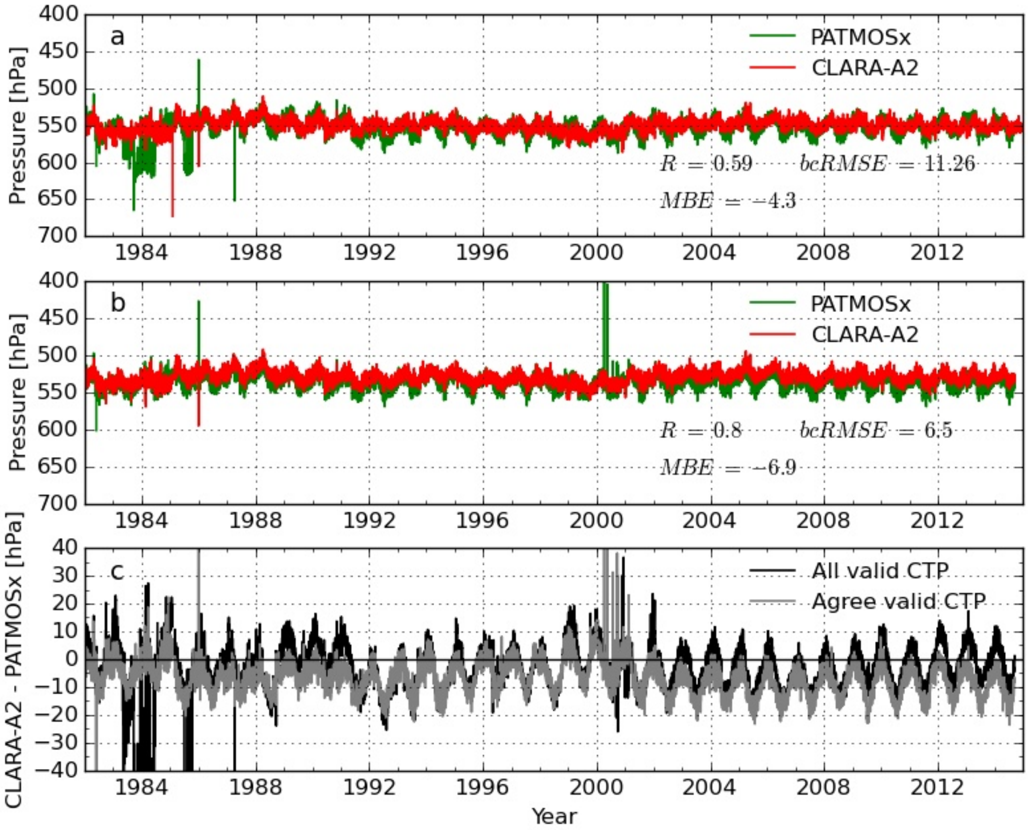


Figure 6.13 *a) Global mean cloud top pressure [hPa] for PATMOS-x (green) and CLARA-A2 (red). Daily averages are computed from all (ascending + descending) satellite overpasses. The R-value (correlation coefficient) is provided in the lower right. Global averages are area-weighted. b) Same as a) but only for pixels both having clouds (common cloud mask)c) Daily global mean cloud top pressure [hPa] difference, defined as CLARA-A2 – PATMOS-x. Results are given both for all cases (black) and for the common cloud mask (grey). The mean bias error, bias-corrected RMS error (both in hPa) and correlation in globally-averaged daily cloud top pressure are provided in the upper right in figures a) and b).*

Table 6.8 *Same as in Table 6.7, but for CTP evaluation from CLARA-A2 level-2b against PATMOS-x level-2b.*

Region / overpass	MBE (hPa)	RMSE (hPa)	r-value
Global / all	-4.3	11.26	0.59
Global / afternoon	-13.41	9.05	0.75
Global / overnight	10.38	9.16	0.72
60°S-60°N / afternoon	-35.56	12.13	0.80
60°S-60°N / overnight	-8.59	10.20	0.77

Accumulated statistics for the full global daily averages, separate overpass nodes (only for observations from satellites with an afternoon/overnight local overpass time), as well as a subset of global CTP excluding high-latitude regions, is presented in Table 6.8. For the first entry in the table (Global/all) corresponding values for the Common cloud mask can be found in Figure 6.13b. The MBE of globally-averaged CTP for CLARA-A2 against PATMOS-x is well within the optimal level-3 product requirement of 80 hPa. MBE and bias-corrected RMSE of CTP are larger for satellites with an afternoon overpass compared to their counterpart overnight overpass. Interestingly, during the afternoon, CLARA-A2 CTP exhibits the largest underestimation, suggesting a potential complication in identifying the cloud top during afternoon convection; this complication may be exacerbated by the relatively large orbital drift in afternoon AVHRR satellites and the influence this drift may have on exceedingly later afternoon observation times. A time series of afternoon 60°S-60°N daily-averaged CTP indicates a rather pronounced trend from 1996 onwards towards higher CTPs for CLARA-A2, while the trend for PATMOS-x is only slightly increasing (not shown). This suggests that orbital drift affecting local observation time may be impacting CLARA-A2 CTPs, but is somehow accounted for by PATMOS-x.

The relative frequency distribution of CTP for each day of January and July 1982-2014 where a valid CTP was retrieved is shown in Figure 6.14. The data are separated into afternoon and overnight overpasses. The distributions of CTP for PATMOS-x indicate a rather distinct bi-modal distribution with a relatively low cloud peak CTP between 800-900 hPa, and a high cloud peak between 150-400 hPa. These features are found for both January (Figure 6.14a) and July (Figure 6.14b). The afternoon and overnight distributions for PATMOS-x are also generally similar, with only modest differences in the maxima peak CTPs (slightly higher cloud tops overnight relative to afternoon). CTP distributions for CLARA-A2 deviate rather dramatically from PATMOS-x. During January, there is a broad CTP distribution with primary peak between 500-700 hPa, which is the pressure range of relative minimum saddle point observed for the PATMOS-x distribution (Figure 6.14a). A relative maximum in the distribution emerges for lower cloud top pressures (higher clouds) during July, broadly consistent with one maxima observed for PATMOS-x (Figure 6.14b). However a secondary peak for mid-level CTPs near 700 hPa still emerges, in the vicinity of where PATMOS-x shows a broad, relative minimum.

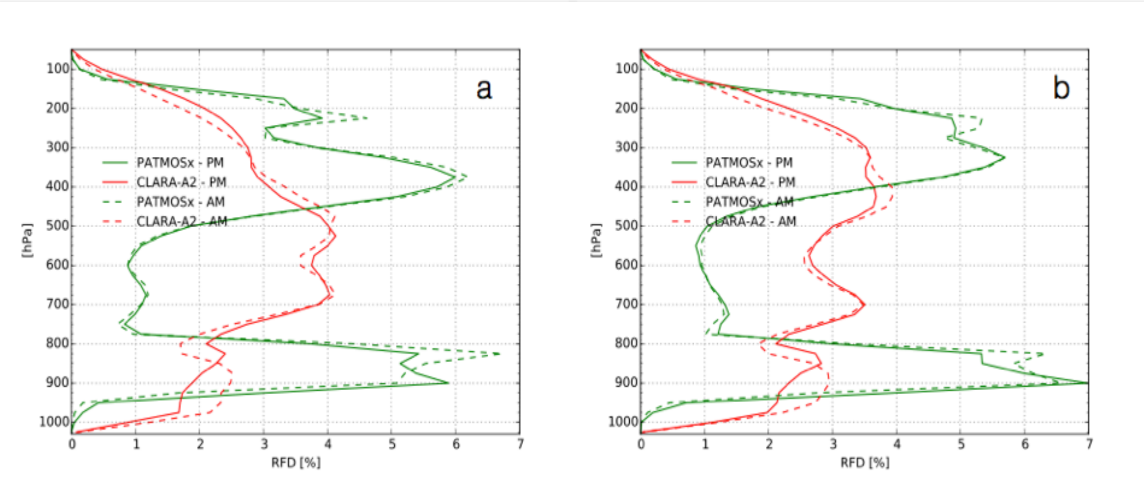


Figure 6.14 *Relative frequency distributions [%] of all valid afternoon (PM, full lines) and overnight (AM, dashed lines) overpasses of level-2b CTP values for CLARA-A2 (red) and PATMOS-x (green) during 1982-2014 for a) January and b) July.*

The distributions shown in Figure 6.14 suggest that CLARA-A2 has an overabundance of a mid-level cloud regime, whereas PATMOS-x tends to classify clouds as either high- or low-level clouds. Both data records indicate a shift towards lower CTPs for overnight nodes. Additionally, cloud tops below ~ 950 hPa are more frequent for CLARA-A2 than for PATMOS-x. Based on the relative distributions in Figure 6.14, it is apparent that the global, daily-averaged CTP shown in Figure 6.13a shows a rather good agreement not because CLARA-A2 and PATMOS-x retrieve similar CTPs, but because the frequency-weighted averages of a broad versus bimodal CTP distribution are more or less similar.

Summary of results

- Very good agreement global, daily-averaged CTP with a mean bias of about -4 hPa relative to PATMOS-x
- Generally CLARA-A2 level-2b CTPs are lower than PATMOS-x, indicating cloud tops that are retrieved slightly higher
- Relative to PATMOS-x, largest CTP underestimates found in Northern Hemisphere summer; smallest in Northern Hemisphere winter
- Afternoon satellite overpasses show a relatively large negative bias, ~ -35 hPa, which is not found for the overnight, ~ -10 hPa
- Bias-corrected RMSEs were consistent, and relatively small, for all sub-regions and overpass nodes examined (bc-RMSE ranging approximately 9 to 12 hPa)
- CLARA-A2 level-2b CTPs show a relative broad distribution, especially for January, compared to PATMOS-x, which distinctly indicates a bimodal frequency distribution dominated by a low-level mode and a high-level mode
- CLARA-A2 July relative frequency distribution indicates a local maxima frequency peak for upper-level clouds, consistent with PATMOS-x. However, CTP distribution still overestimating mid-level frequencies and missing low-level local maxima

6.1.2.3 Evaluation of CTH level-2b products against PATMOS-x for July 2008

The CLARA-A2 CTH, LWP and IWP level-2b products were compared with PATMOS-x level-2b for the month July 2008. Retrievals from the afternoon satellite NOAA-18 were analysed. In addition, the morning satellite NOAA-17 was considered, on which AVHRR channel 3a was active during daytime rather than channel 3b. The comparisons were restricted to daytime (here defined by a solar zenith angle smaller than 82 degrees), i.e. ascending orbits for NOAA-18 and descending orbits for NOAA-17. The CLARA-A2 data were subsampled to the PATMOS-x spatial resolution of 0.1 x 0.1 degrees, and only grid cells classified as cloudy by both data records were included in the comparisons. For the liquid/ice cloud property analyses only grid cells classified as that particular phase by both data records were included. Aggregation over time, to generate global monthly maps, was done by linear averaging of the properties from all cloud / liquid-phase cloud / ice-phase cloud occurrences for a particular grid cell in the month.

A specific study focussing on one selected month (July 2008) was also carried out for the cloud top height (CTH) product. Figure 6.15 shows that the spatial distributions of this parameter are quite similar, with highest clouds occurring in the Tropics and lowest clouds in the marine stratocumulus areas. A clear difference is apparent for the latter though, with CLARA-A2 placing these clouds higher than PATMOS-x.

Figure 6.16 further illustrates the differences for low clouds. While PATMOS-x CTH has a relatively narrow peak between 0.5 and 2 km, CLARA-A2 shows a much broader distribution between the surface and 4-5 km. For higher clouds the correspondence between the data records is much better.

NOAA-17 results are very similar to NOAA-18 and are therefore not shown.

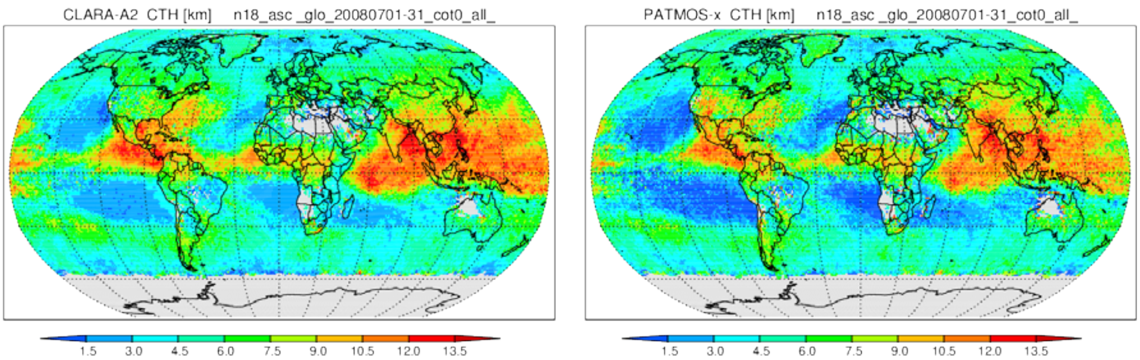


Figure 6.15 *Mean daytime cloud top height in km from NOAA-18 for July 2008. Left: CLARA-A2; right: PATMOS-x. Grey areas indicate no data because no clouds were detected or the solar zenith angle was too high during the entire month*

.

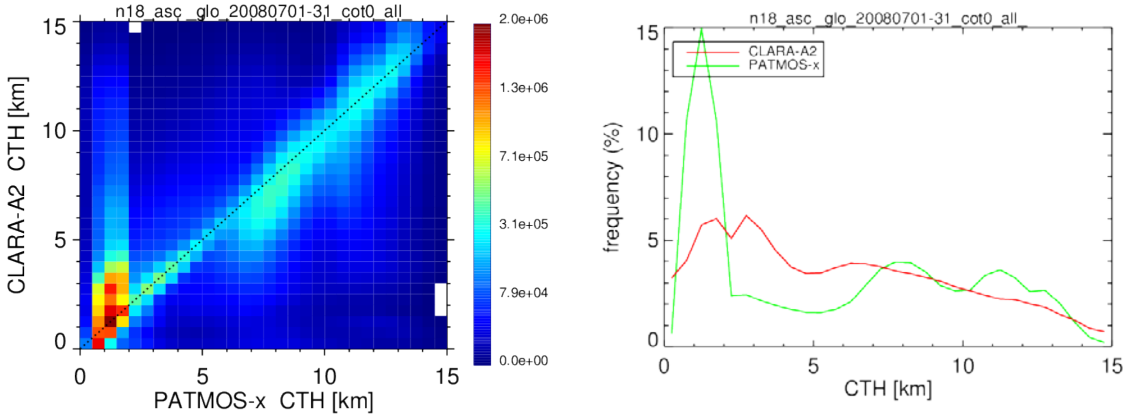


Figure 6.16 *Pixel-level comparison between CLARA-A2 and PATMOS-x daytime CTH from NOAA-18 for July 2008. Left: scatter-density plot in which the colours indicate the number of pixels (level-2b grid cells) with the particular CLARA and PATMOS CTH values; right: 1-dimensional histograms.*

6.1.2.4 Evaluation of LWP level-2b products against PATMOS-x for July 2008

For the evaluation of liquid water path we focus on the two directly retrieved parameters cloud optical thickness (COT) and droplet effective radius (REFF) which together determine LWP. Figure 6.17 shows global maps of these cloudy-sky averaged properties. COT from CLARA-A2 and PATMOS-x is very similar with significant differences only occurring in the Arctic, where PATMOS-x has higher values. Both data records yield very large COT over Greenland, which can be attributed to retrieval problems over the bright ice-covered surface. The spatial distributions of REFF are also similar, with generally lower values over land than over sea.

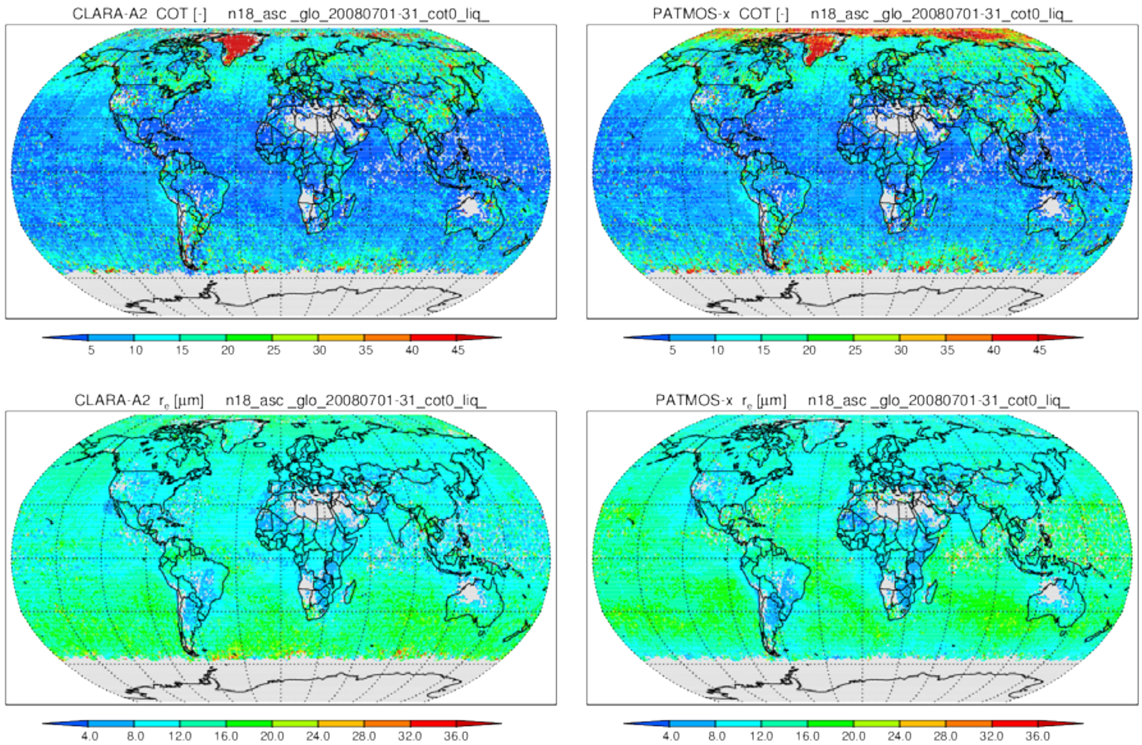


Figure 6.17 *Mean liquid cloud optical thickness (top) and effective radius from NOAA-18 for July 2008. Left: CLARA-A2; right: PATMOS-x. Grey areas indicate no data because no clouds were detected or the solar zenith angle was too high during the entire month.*

Pixel-based comparisons are presented in Figure 6.18. This confirms that COT retrievals are in very good agreement except for thin clouds: while CLARA-A2 has values down to 0.1 (the minimum retrieved), PATMOS-x yields hardly any COT below 1. Effective radii are also well correlated, but there are some peculiar differences: (i) CLARA REFF is overall 1-2 µm larger, and (ii) for thin clouds CLARA REFF is weighted with a climatological value of 8 µm, yielding a peak at that value. The combined effect of COT and REFF explains the picture for LWP. In particular, CLARA has a much higher occurrence frequency of LWP < 10 g m-2, while PATMOS-x has more clouds with 10 g m-2 < LWP < 50 g m-2.

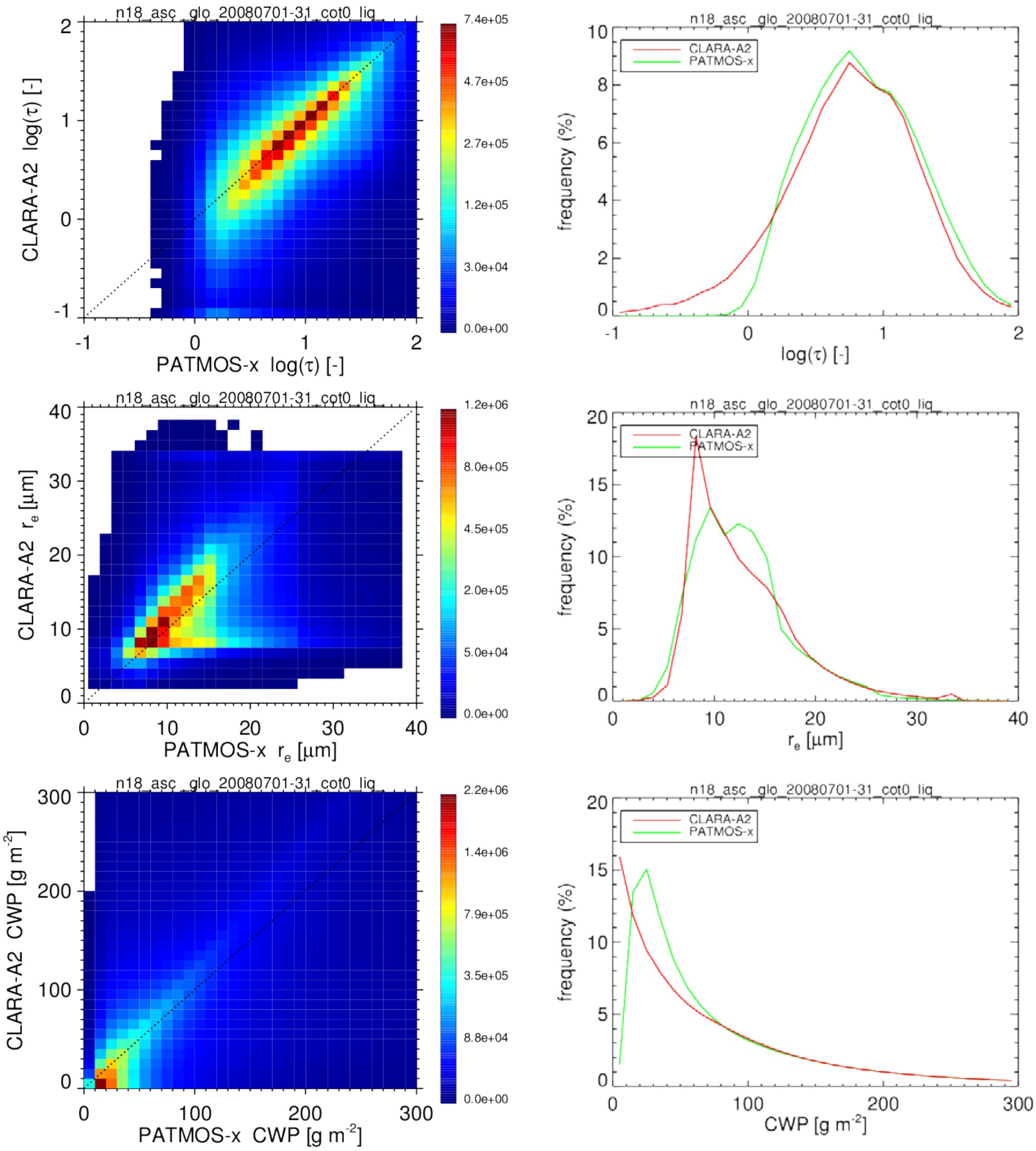


Figure 6.18 *Pixel-level comparison between CLARA-A2 and PATMOS-x liquid COT (top: note that the logarithm of COT is shown), liquid REFF (middle), and LWP (bottom) from NOAA-18 for July 2008. Left: scatter-density plots in which the colours indicate the number of pixels (level-2b grid cells) with the particular CLARA and PATMOS parameter values; right: 1-dimensional histograms.*

The comparison of NOAA-17 retrievals yields similar results regarding liquid COT (not shown), but very different results regarding liquid REFF (Figures 6.19 and 6.20). The bulk of the REFF retrievals are in better agreement and both histograms are broader than for NOAA-18. For CLARA-A2 a distinct peak at the maximum retrieved value of 34 µm is observed, while PATMOS-x yields even considerably larger REFF. These differences must be related to the use of channel 3a on NOAA-17 rather than channel 3b on NOAA-18. Even if there are clear physical reasons for REFF retrievals based on channels 3a or 3b being different, this is a feature that has to be kept in mind when using the respective REFF (and LWP/IWP) products.

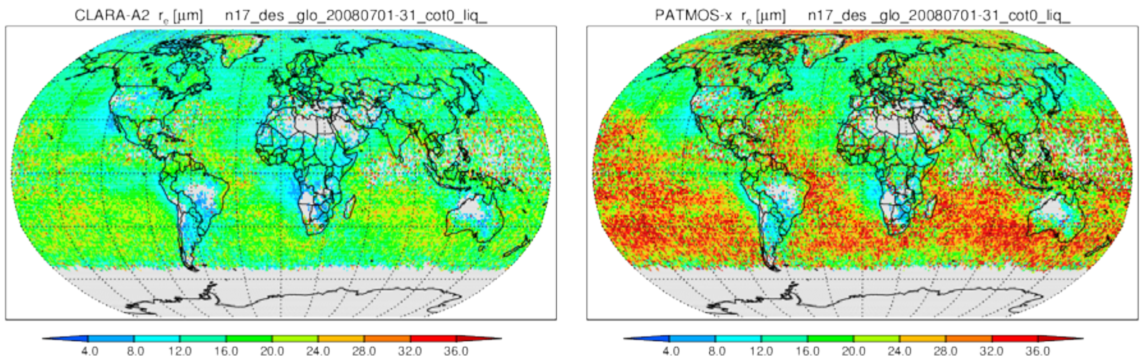


Figure 6.19 *As Figure 6.17, but now for NOAA-17, and only liquid REFF is shown.*

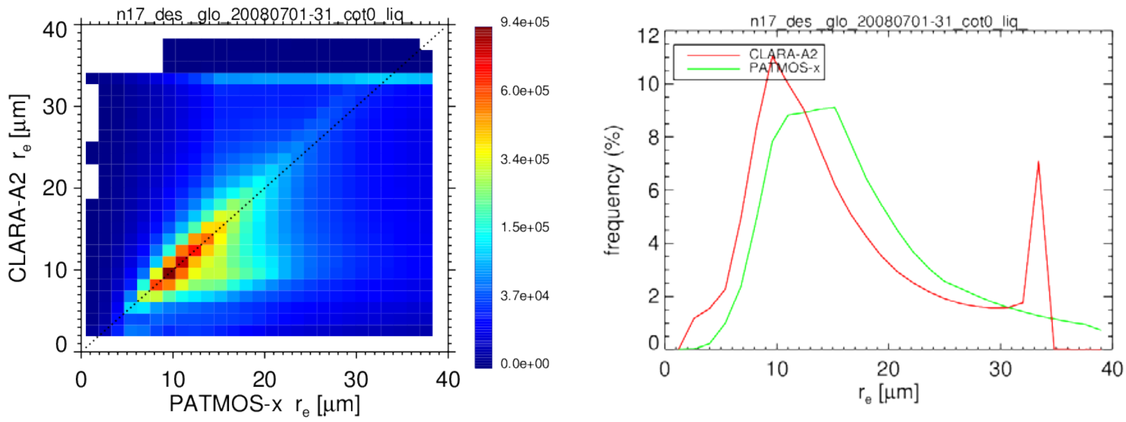


Figure 6.20 *As Figure 6.18, but now for NOAA-17, and only liquid REFF is shown.*

6.1.2.5 Evaluation of IWP level-2b products against PATMOS-x for July 2008

CLARA-A2 and PATMOS-x ice cloud properties are compared in Figures 6.21 and 6.22. The global maps (Figure 6.21) show very good agreement in COT with similar features as observed for liquid water clouds. Differences in REFF are larger, with in particular overall higher values for PATMOS-x. The scatter plots and histograms (Figure 6.22) appear to show two regimes for REFF: (i) values smaller than 10 µm, for which the agreement is quite good, and (ii) values between 15 and 30 µm, for which CLARA-A2 REFF is about 5 µm smaller than PATMOS-x. These differences may be related to the ice models used for the single scattering calculations: imperfect hexagonal ice crystals for CLARA-A2 vs. roughened aggregates for PATMOS-x (Baum et al., 2012). Despite these differences in REFF, the IWP histograms are in relatively good agreement.

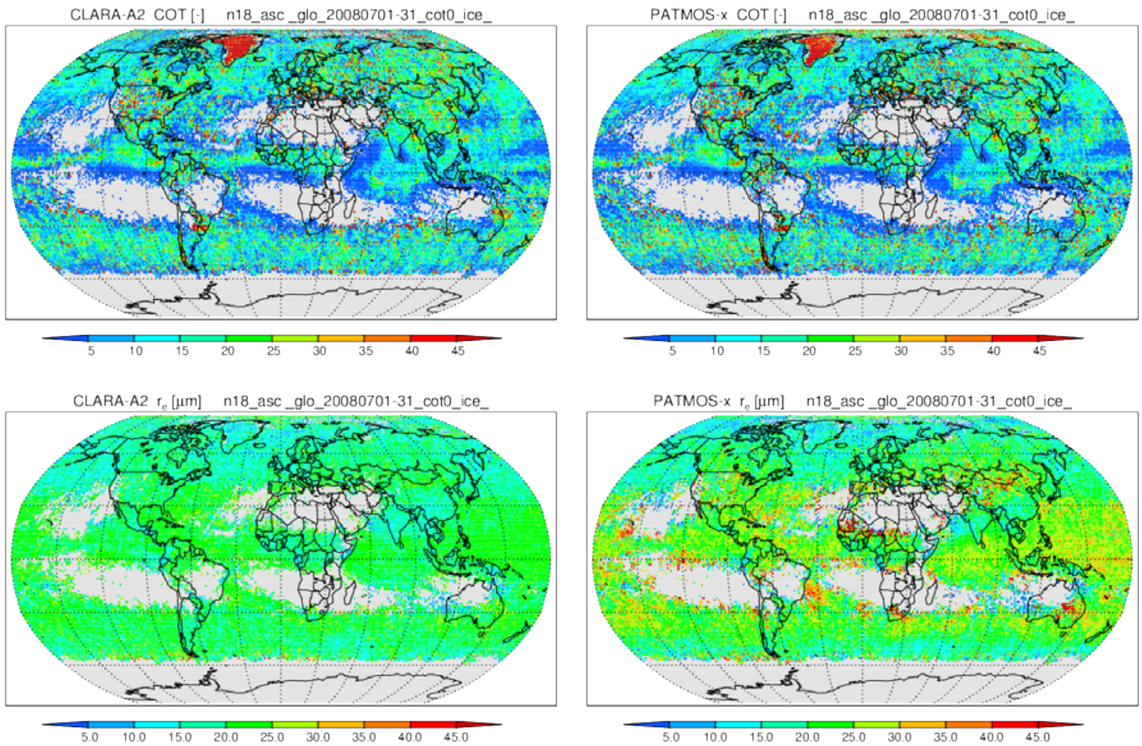


Figure 6.21 *As Figure 6.17, but now for ice cloud properties.*

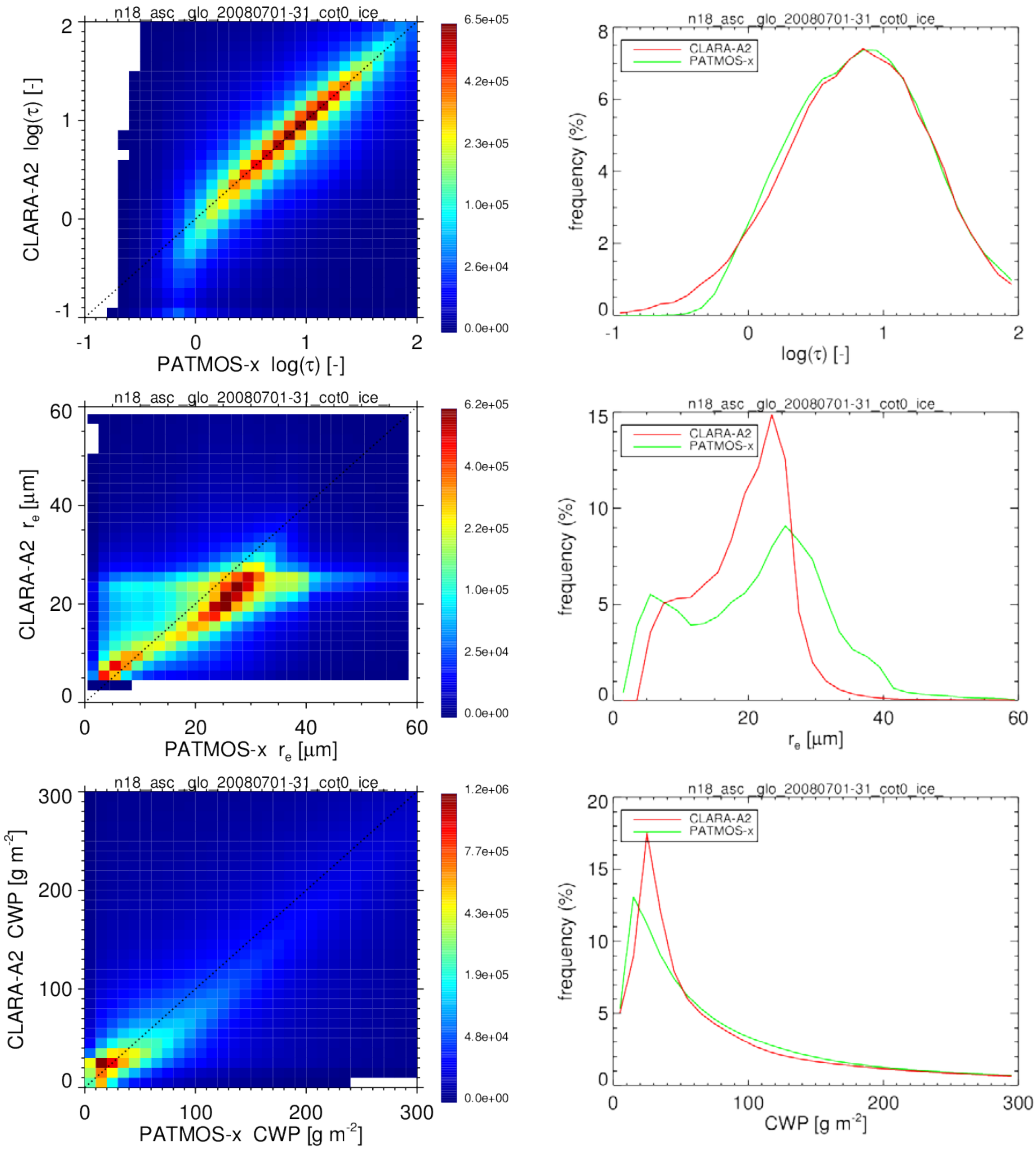


Figure 6.22 *As Figure 6.21, but now for ice cloud properties.*

Like liquid REFF, the ice REFF retrievals also depend strongly on the shortwave-infrared channel used. Thus, for NOAA-17 the results are quite different than for NOAA-18. In particular, PATMOS-x yields much larger particle sizes (Figure 6.23). The scatter plot of pixel-based ice REFF (Figure 6.24, left panel) is qualitatively similar to the corresponding plot for NOAA-18 (Figure 6.20), with an even somewhat larger difference for effective radii above 15 µm. The histograms (Figure 6.22, right panel) show that PATMOS-x REFF extends to values of about 60 µm rather than the 40 µm for NOAA-18. In comparison, CLARA-A2 has relatively smaller differences between the NOAA-17 and NOAA-18 histograms.

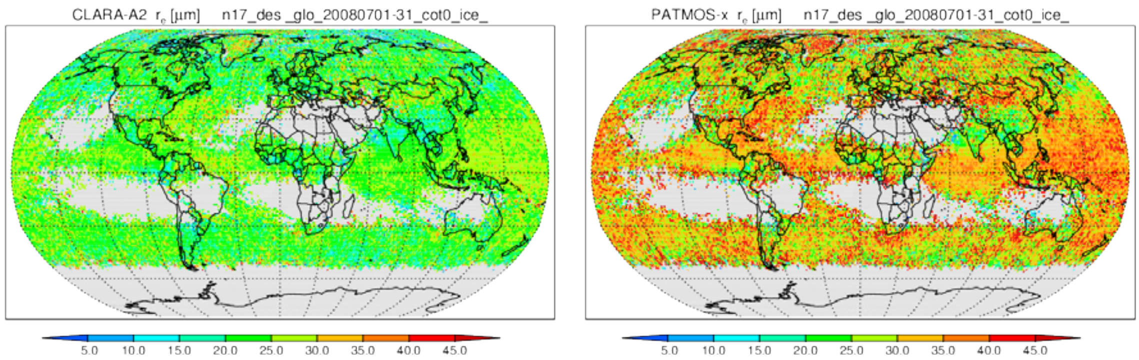


Figure 6.23 *As Figure 6.21, but now for NOAA-17, and only ice REFF is shown.*

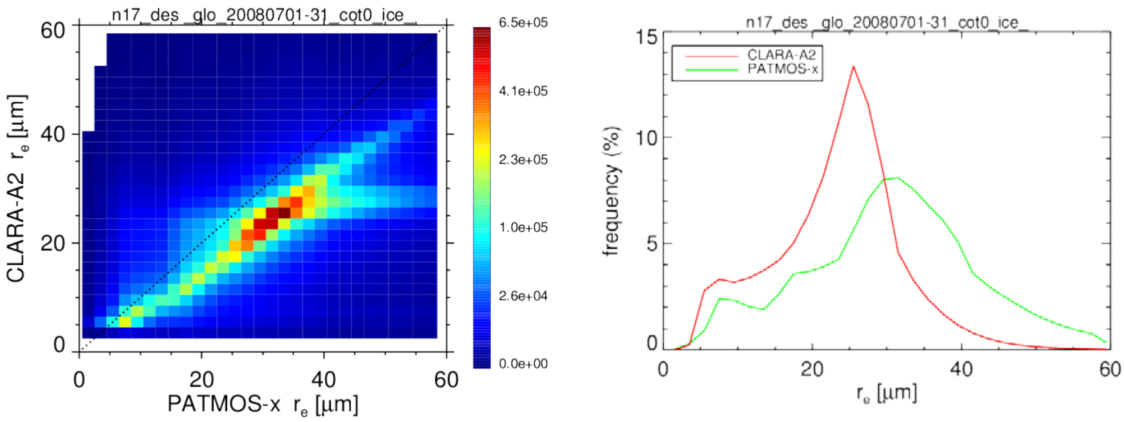


Figure 6.24 *As Figure 6.22, but now for NOAA-17, and only ice REFF is shown.*

6.1.3 Evaluation against DARDAR (Cloudsat-CALIPSO)

6.1.3.1 Evaluation of CPH against DARDAR

The retrieved CLARA-A2 AVHRR level-2 cloud phase is compared to DARDAR retrievals for the month of January 2008. All DARDAR profiles were checked for the number of different cloud phases within the profile. Only those profiles which consist of a single cloud phase (either liquid or ice) have been taken into account. Collocations between the A-Train (DARDAR product) and NOAA-18, NOAA-17 and METOP-A have been considered. While NOAA-18 yields collocations at all latitudes, NOAA-17 and METOP-A only have overlap close to the poles. These regions are characterized by difficult retrieval situations with frequent occurrence of supercooled liquid layers on top of ice clouds and frequent low-altitude clouds over highly reflective surfaces. Henceforth, this evaluation has been restricted to the NOAA18 satellite for latitudes between -75 and 75 degrees. The results presented below are based on the measurements in January 2008. The same comparisons have been made for the month of July of the same year, resulting in very similar results, verifying that the presented results are robust.

In the cloud phase comparison only those profiles are taken into account for which DARDAR shows a single cloud phase, e.g., supercooled layers over rain or liquid clouds are taken into account but profiles with ice layers over liquid clouds have been removed from the comparison. The cloud phase verification scores are shown in Figure 6.25. The results show an increasing probability of detection of ice clouds with the optical thickness at which the phase in the DARDAR profile is probed. This is not because the phase in the DARDAR profile changes (as mentioned before only single-phase DARDAR profiles are considered) but because thin clouds are removed from the sample when going to the right in Figure 6.25. These thin clouds tend to be ice clouds, sometimes erroneously labelled liquid in CLARA-A2. Similar to the increase in POD ice, the false alarm ratio decreases at higher optical thicknesses. Overall, the skill scores are significantly higher than found in the CLAAS-2 (SEVIRI-based) cloud phase evaluation [RD 8]. The main reason for this is thought to be the lower viewing angles for the polar orbiter collocations with DARDAR compared to the geostationary satellite collocations with DARDAR.

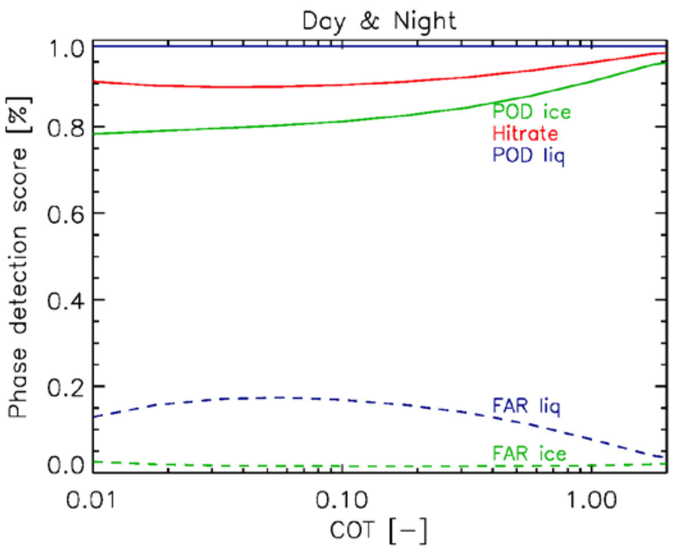


Figure 6.25 *CLARA-A2 cloud phase hit rate, probability of detections and false alarm ratios for both liquid and ice clouds as a function of the integrated optical thickness from the top of the cloud. Results are for NOAA-18 collocations in January 2008. Only single cloud phase DARDAR columns were used in these statistics and both day and night observations were taken into account.*

6.1.3.2 Evaluation of IWP against DARDAR

The CLARA-A2 IWP based on the AVHRR measurements is retrieved using the CPP algorithm. Within CPP, IWP is a secondary product, based on the direct retrievals of effective radius and optical thickness. Before looking at the IWP, statistics of the two direct retrievals are analysed. In Figure 6.26, the CLARA-A2 vs. DARDAR single-layer ice cloud optical depth comparison is shown. The distribution contours show the number of points enclosed, i.e. the black area shows the top 20% of the number of points, this part of the distribution is correlated (0.78 in log space) and lies along the one-to-one line. The correlation of the two data sets drops to 0.63 (in log space and 0.4 in linear space) when comparing all available points.

The second direct retrieval from the CPP algorithm used to calculate the CLARA-A2 results is the effective radius. The retrieval uses the Nakajima and King (1990) approach, deriving both optical depth and effective radius simultaneously using pre-calculated lookup tables (LUTs). The DARDAR product provides an effective radius profile and not the radiative layer effective radius observed by the AVHRR imager. To enable the comparison, the layer averaged effective radius from DARDAR, $\hat{R}_{\text{eff}}$, was calculated in three ways as a function of the integrated optical thickness from the top τ^*:

$$(1)\ \hat{R}_{\text{eff}}(\tau^*) = R_{\text{eff}}(z(\tau^*))$$

$$(2)\ \hat{R}_{\text{eff}}(\tau^*) = \frac{\int_{z(\tau=0)}^{z(\tau=\tau^*)} R_{\text{eff}}(z)\, dz}{\int_{z(\tau=0)}^{z(\tau=\tau^*)} dz}$$

$$(3)\ \hat{R}_{\text{eff}}(\tau^*) = \frac{\int_{z(\tau=0)}^{z(\tau=\tau^*)} R_{\text{eff}}(z)\, \alpha(z)\, e^{-\tau(z)}\, dz}{\int_{z(\tau=0)}^{z(\tau=\tau^*)} \alpha(z)\, e^{-\tau(z)}\, dz}$$

The first method simply picks the effective radius at τ^*. The second method is a plain average over the upper part of the cloud, while the third method is a weighted average taking into account the extinction and transmission (α). Note that τ^* is maximized at the total optical thickness of the cloud.

In Figure 6.27 the results are shown using method 3 and for τ^*=1. This procedure weighs the R_{eff} towards cloud top, resulting in a slightly lower R_{eff} value in comparison to the plainly averaged R_{eff} (method 2) and a lot smaller in comparison to the local effective radius at an optical depth of 1 method 1, not shown). Even with this focus on the upper part of the ice clouds the resulting DARDAR $\hat{R}_{\text{eff}}$ distribution is a lot wider (between 20 and 80 microns) than for CLARA-A2. When looking deeper into the cloud ($\tau^* > 1$), the DARDAR R_{eff} distribution moves to larger sizes and vice versa. The CLARA-A2 effective radius distribution in contrast is very narrow and peaks between 20 and 30 microns, resulting in the end in no correlation between the two distributions.

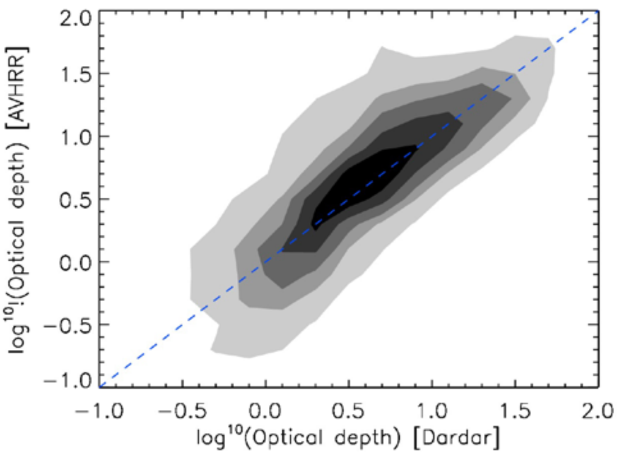


Figure 6.26 *Ice cloud optical thickness distribution comparing the DARDAR and CLARA-2 retrieved collocated values. The blue dashed line shows the 1-1 line with the greyscales indicating the regions enclosing the 20, 40, 60, 75, and 90% of points with the highest occurrence frequency.*

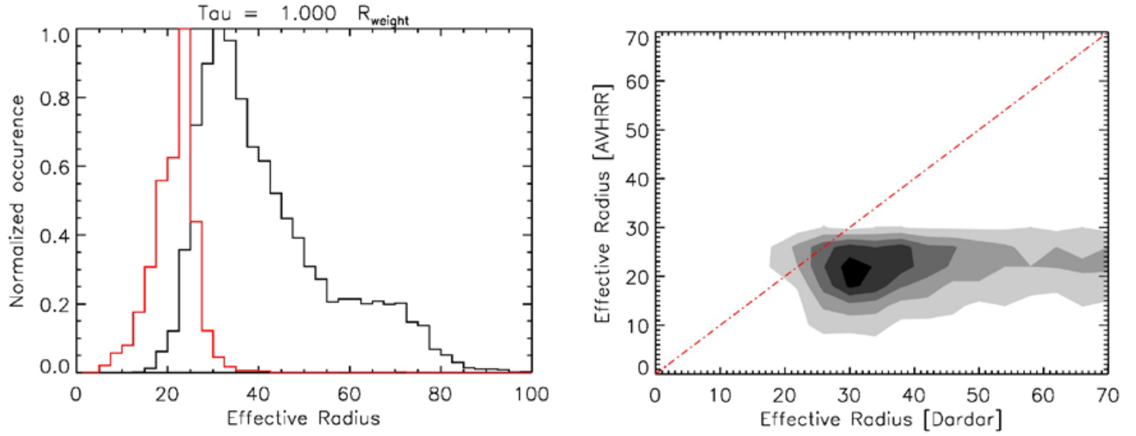


Figure 6.27 *Comparison of CLARA-A2 ice effective radius and DARDAR weighted effective radius from cloud top to an optical depth of 1 (or to cloud base if the total optical depth is smaller than 1). The left plot shows 1D-histograms with CLARA-A2 indicated in red and DARDAR in black; on the right a scatter density plot is shown. The dynamic range of the DARDAR retrievals is a lot larger resulting in no correlation between the two distributions. The greyscales indicate regions enclosing the 10, 30, 50, 70, and 90% of points with the highest occurrence frequency.*

The ice water path (IWP) is proportional to the product of the two parameters discussed above. Due to the differences seen in the effective radius distributions the overlay along the 1-1 line for the optical depth data is converted into a curved 2D-occurrence distribution (Figure 6.28, left panel), with overall lower IWP values for the AVHRR retrievals in comparison to the DARDAR retrievals. The occurrence distributions (dynamic range) of the individual data sets however are very similar (right panel).

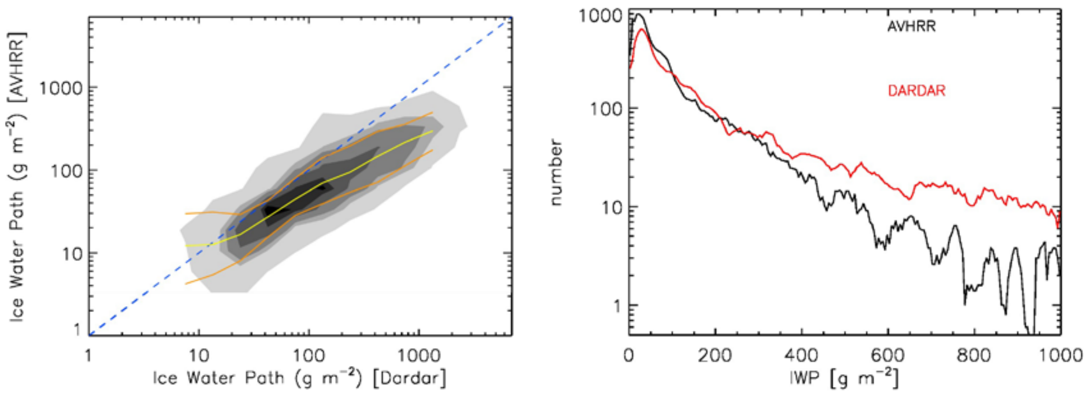


Figure 6.28 *Left panel: CLARA-A2 IWP vs. DARDAR IWP. The yellow line depicts the median and orange the 16th/84th percentiles of the CLARA-A2 distribution at the local DARDAR IWP. Right panel: 1D-histograms of DARDAR and CLARA-A2 IWP for the same collocations. The greyscales indicate regions enclosing the 10, 20, 40, 60, and 75% of points with the highest occurrence frequency.*

The distributions look similar to the ones presented in Eliasson et al. (2013), with the curve for low IWP values as seen in their Figure 6 for MODIS and PATMOS-x. When all observations are included the IWP bias is 153.1 and the bias-corrected RMS is 366.8 Focussing on the black region with the 10% highest-density points, the bias drops to 29.4 and the bias-corrected RMS to 21.5 The results once more show that a comparison of passive versus active instruments (and between different instruments in general) is tricky due to the different microphysical assumptions and the difference between profile information vs. column averaged (but weighted to the top of the cloud) measurements. Both influence the effective cloud depth, from where most of the information comes.

There is a significant improvement in the optical depth retrieval and cloud phase determination in comparison to a geostationary instrument, where the latter suffers from large differences in viewing angle and therefore collocation problems. What is clear is that the CLARA-A2 retrieved effective radius is on the small side, even though it is radiatively internally consistent with the CLARA-A2 microphysical assumptions. For CLARA-A3 it will be considered to alter the microphysical assumptions from the roughened randomly oriented hexagons (Hess et al., 1998) in CLARA-A2 to a more general aggregate habit mode, e.g., Baum et al (2012) or Baran et al. (2005). This may enable the retrieval of a more consistent IWP with respect to the active instruments in a future data record.

6.2 Evaluation of AVHRR level-3 products (including joint histograms)

This section covers the evaluation of CLARA-A2 level-3 products. These consist of daily and monthly aggregations. The evaluation is organized according to Table 6.9.Notice that section 6.2.5 contains validation of the entire group of CPP products using several references.

Table 6.9 *Overview of reference data records used for the evaluation of CLARA-A2 level-3 parameters.*

Section	Reference observations	Parameters
6.2.1	SYNOP	CFC
6.2.2	MODIS	CFC,CTP
6.2.3	ISCCP	CFC,CTP
6.2.4	PATMOS-x	CFC, CTP
6.2.5	PATMOS-x, MODIS, ISCCP, UWisc	CPP (CPH, LWP, IWP)
6.2.6	MODIS	JCH
6.2.7	MODIS, PATMOS-x	Process-oriented studies of CFC, LWP, IWP

6.2.1 Evaluation of CLARA-A2 CFC level-3 with SYNOP

SYNOP total cloud cover observations are used for the evaluation of level-3 cloud cover estimates. For the level-3 comparison the available number of SYNOP monthly mean estimations reflects the known geographically unbalanced distribution of the synoptic stations: the majority of the stations are located in the northern mid-latitudes while there are fewer stations over large parts of Africa and the northern part of South America. This uneven distribution has to be kept in mind when looking at accumulated statistics. Also the number of available SYNOP stations increases with time. To account for this effect only stations are included into this analysis, that cover more than 95% of the full time period.
In order to evaluate the performance of the CLARA-A2 cloud fraction both mean error (accuracy parameter; bias) and bias-corrected Root Mean Square errors (precision parameter; bc-RMSE) have been calculated and then compared to the defined target requirements as specified in Table 4.1.
For the level-3 comparison the CLARA-A2 monthly mean product generated from all the satellites available was compared against SYNOP monthly mean cloud cover calculated based on daily means. Only those stations and months were taken into account where at least 6 observations per day at 20 days of the respective month are available.

Figure 6.29 shows the time series of the monthly mean global fractional cloud cover. The black curve shows the SYNOP, the red CLARA A2, and grey shows the former CLARA data sat. The time series of CLARA A2 monthly means is based on all available satellites, which are aggregated to one monthly mean. Each satellite has different orbit and overpassing times which enables a better representation of the diurnal cycle in the satellite data record. But, the number of available satellites is not stable over the entire time series, which causes less representative results when fewer satellites are included. However, using all available satellites will make the data record as good as possible in terms of comparability to full day SYNOP observations.

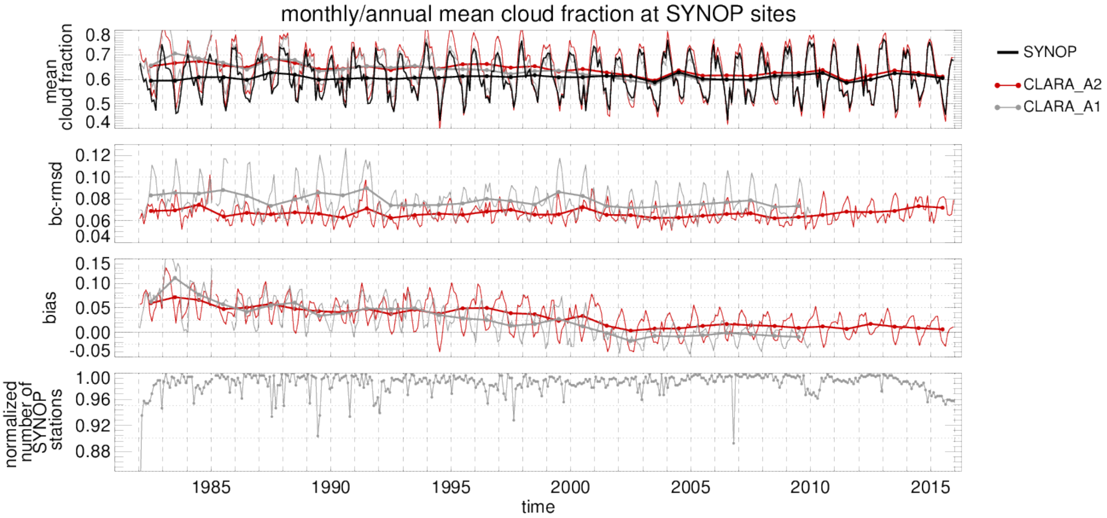


Figure 6.29 *Time series of mean cloud cover for CLARA-A2 (red), CLARA-A1 (grey), and SYNOP (black) (upper panel), bias-corrected RMSE (second panel), bias (third panel), and the number of stations (lower panel) normalized to 1 for the entire period 1982-2015.*

In the first years until 1987 the overestimation of cloud cover of the CLARA A2 data record has been decreased compared to CLARA-A1. This is related to removal of false clouds over semi-arid regions and the improved handling of noise in channel 3b of the AVHRR-2 instruments.

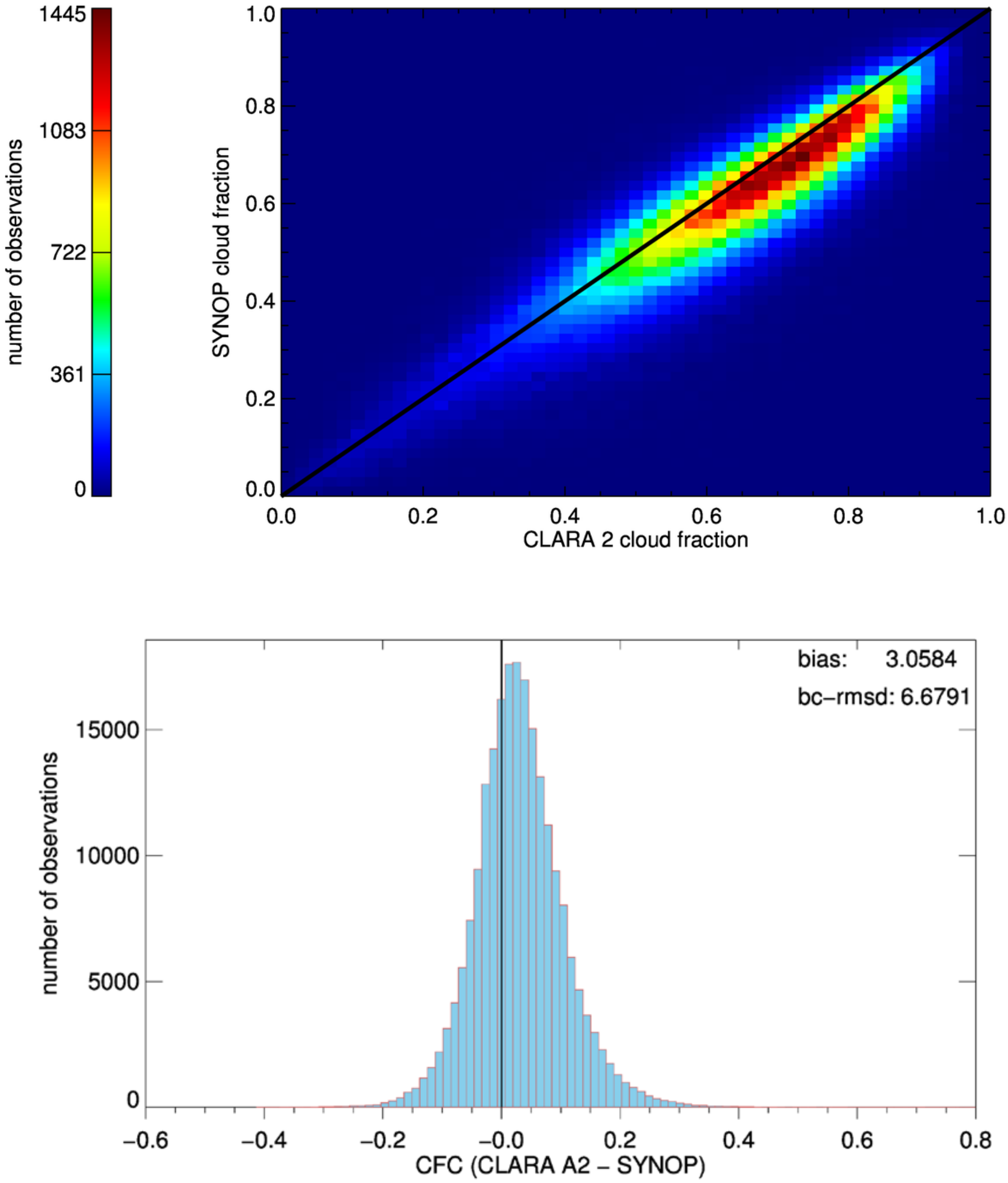


Figure 6.30 *2D-scatter plot of the monthly mean cloud cover shown by CLARA A2 and SYNOP (top) and the histogram of the difference between CLARA A2 and SYNOP (bottom) for the entire period.*

The bc-RMSE shows clearly reduced values for the new data record. This means an improved precision of the new data record. Also the bc-RMSE show a very smooth time series with low variation. Only in the monthly means, larger seasonal cycles are observed for the years until NOAA12 and the first 2 years of NOAA15. The bias, shown in the third panel, is continuously decreasing with the time. But this decrease could be reduced, compared to the previous version. In the new version, the bias is lower in the first years and positive over the full timeseries. The decrease of the bias has basically one reason. The CLARA-A2 data record uses an increasing number of satellites with time. From 1991 the morning orbit NOAA-satellites was included in CLARA-A2 and from 1999 the NOAA15-satellite (the first of the

NOAA-KLMN series of satellites) with a revised AVHRR instrument (AVHRR/3) was introduced. The increased lifetimes of the NOAA satellites (in particular NOAA-12 and NOAA-14 and the KLMN-satellites) lead consequently to the situation, that AVHRR observations from at least 3 satellites (in 2009 up to 6 satellites – see Figure 3) have been available simultaneously during the last 15 years of the time series. This increased the ability of representing the diurnal cycle of the fractional cloud cover, which can be well identified by the decrease in the bias in the year 2001.

Figure 6.30 shows a more detailed analysis of the validation of CLARA-A2 monthly means against SYNOPs. The upper panel shows the 2-dimensional histogram comparing the two data records. Here, a small overestimation of CLARA A2 is found between 0.4 and 0.9 fractional cloud cover. For lower values results are well distributed between low over- and underestimation with low scattering. Generally, the histogram shows good agreement between both data sets with only minor scattering. The bottom panel presents the distribution of the differences (CLARA-A2 - SYNOP). Nearly all differences are within +- 0.2 fractional cloud cover and the curve shows no significant skewness or kurtosis. The peak is shifted towards the right by 0.026 (more or less identical to the average bias).

Summary of results:

- Good agreement in general: the bias lies mostly within +/- 20% cloud amount (~ 2 octa)
- After 2001 the data record shows a very stable and low bias.
- Overall, the variability is very low and stable.
- The overall mean error remains stable over time and lies at or within the target accuracy of +/- 5 % cloud amount (exception only in 1983)
-

Table 6.10 *Compliance matrix of found global CFC monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against SYNOP observations.*

	CFC product requirements level-3 (MM)			SYNOP level-3(1982-2015)
	Threshold	**Target**	**Optimal**	
Bias	10 %	5 %	2 %	3.1 %
bc-RMS	20 %	10 %	5 %	6.7 %

6.2.2 Evaluation against MODIS

6.2.2.1 Evaluation of CLARA-A2 CFC level-3 products

In this section CFC level-3 products (monthly means) of CLARA-A2 are compared to MODIS (MOD08_M3) equivalents. The comparison is based on the entire available time series. For the AQUA satellite this is 2002-2015 and for TERRA 2000-2015. For comparing CLARA-A2 against AQUA only afternoon satellites (NOAA16, NOAA18, NOAA19) have been used. In the case of TERRA only morning satellites have been used (NOAA15, NOAA17, METOP A, METOP B). For both, only prime satellites have been considered. Prime satellites are the satellites being closest to the nominal morning and afternoon orbits (in practice, the satellites with shortest time since launch and thus being exposed to minimum orbital drift). Results are shown in Figure 6.27 exemplarily for the analysis of CLARA-A2 against MODIS/AQUA. The results against TERRA are similar except for the years 2014 and 2015, when TERRA shows an increasing trend. For all comparisons of this kind, both data records are compared month for month. For each month the spatial resolution of the global grid is reduced to the lower resolved grid. Here, this is the used MODIS dataset with a spatial resolution of 1°. Only if the spatial resolution is identical, two datasets can be compared thoroughly. . Then the global mean, bias and bc-RMSE is computed based on each grid box, that has valid information in both data records. Finally, these results are put together to the time series. Based on all biases and bc-RMSEs the final bias and bc-RMSE of the available time series is calculated.

From the first two panels in Figure 6.31 a negative bias of CLARA-A2 can be seen with a stable bc-RMSE below 10 %. Also the spatial distribution of the difference shows a general underestimation of the cloud cover except for the desert region in north and South Africa, Arabia, Australia, and over Tibet. Regions with very strong convection, such as the ITCZ or India show nearly no bias. The four panels in the bottom show the underestimation by CLARA-A2 as well. Also the relatively low variation, indicated by the small bc-RMSE in the time series plot, is reflected in the 2D histogram. The averaged zonal mean plot shows that the underestimation by CLARA-A2 is well distributed over all latitudes with a minimum between 60°S and 80°S.
In Table 6.11 the bias and bc-RMSE are shown for the MODIS comparison. Here, also the comparison against MODIS TERRA is considered, showing comparable results. The slightly larger bc-RMSE and bias can possibly be explained by the noticed increase of TERRA cloud fraction after about 2009.

Summary of results:

- Good agreement in general cloud pattern descriptions but overall lower CFC values for CM SAF (about -5 %)
- Very good results in terms of stability of the bias and bc-RMSE (further discussed in section 6.3.2). MODIS data are supposed to give a better cloud detection capability which is also indicated by the stable negative bias. Nevertheless, the small bc-RMS indicates a very good agreement with MODIS data.
- Positive deviations are found exclusively over desert areas.
- ITCZ and subpolar oceans (off the coast of Antarctica and north of the coast of Russia and Canada) show nearly no deviation.
- Noticeable negative deviations are also seen over oceanic areas, especially in the subtropical areas outside the stratocumulus regions.

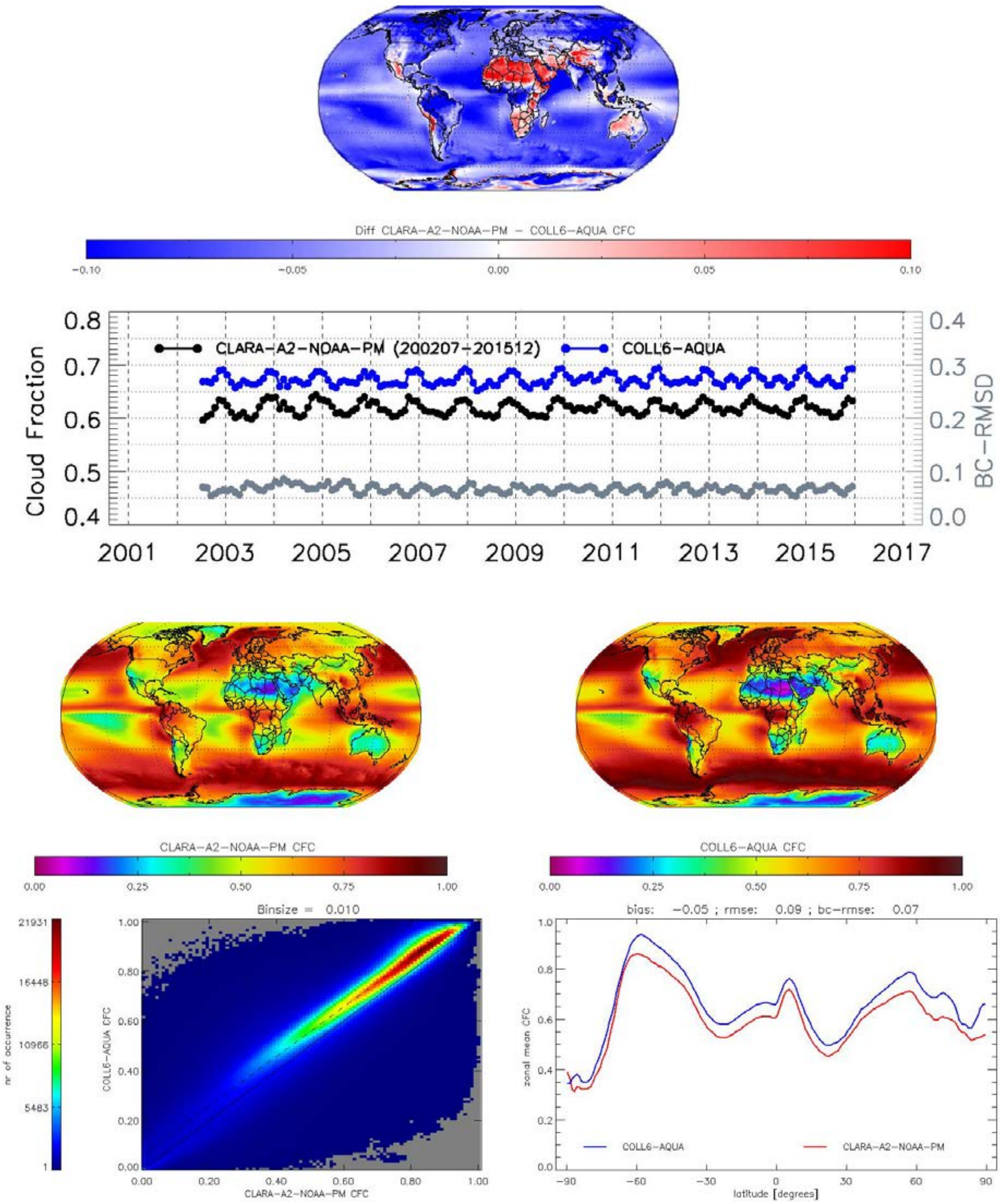


Figure 6.31 *This figure shows the cloud cover comparison of CLARA-A2 afternoon satellites and MODIS collection 6 AQUA monthly means for the entire available time series 2002-2015. The top panel shows the difference plot, the panel below the time series and the bc-RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA-A2 left and MODIS right). The bottom left panel shows the 2D histogram of all data points in time and space and the bottom right panel the averaged zonal mean for CLARA-A2 in red and MODIS Aqua in blue.*

Table 6.11 *Compliance matrix of found global CFC monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against MODIS results (consistency check).*

	CFC product requirements level-3 (MM)			MODIS/Aqua (2002-2015)	MODIS/Terra (2000-2015)
	Threshold	Target	Optimal		
Bias	10 %	5 %	2 %	-5.4 %	-6.4 %
bc-rms	20 %	10 %	5 %	6.8 %	7.9 %

6.2.2.2 Evaluation of CLARA-A2 CTP level-3 products

In this section CTP level-3 (monthly means) products of CLARA-A2 are compared to MODIS (MOD08_M3) equivalents. The comparison is based on the full available time series. For MODIS onboard of the AQUA satellite this is 2002-2015 and for TERRA 2000-2015. For the comparison of CLARA-A2 against AQUA only afternoon satellites (NOAA16, NOAA18, NOAA19) have been used. In the case of TERRA only morning satellites have been used (NOAA15, NOAA17, METOP A, METOP B). For both, only prime satellites have been considered. Results are shown in Figure 6.32 exemplarily for CLARA-A2 against AQUA. The results against TERRA look similar.

From the first two panels a negative bias of CLARA-A2 can be seen with a stable bc-RMSE at about 60 hPa. The distribution of the difference shows a general underestimations of the CLARA A2 cloud top pressure, i.e., clouds are seen at a higher altitude in CLARA-A2. This is stronger over the subtropical oceans off the west coasts of continents. This is most likely connected to problems in correctly identifying the temperature inversion above the maritime boundary layer. Over the deserts in North Africa and Saudi Arabia the cloud top pressure is overestimated by CLARA-A2. For the inner tropics CLARA A2 also shows significant underestimation of the cloud top pressure.. only in the central southern subtropical Atlantic and Pacific ocean and the mid latitudes the bias is nearly balanced..

The four panels in the bottom show the underestimation of CLARA-A2 CTP as well. The global maps of the averaged cloud top pressure show a good agreement of the general patterns. But it also clearly shows that the low clouds are identified by CLARA-A2 too high in the subtropics. In the 2D histogram (bottom left) all data points are close to the one-by-one line, but the scatter is much stronger, than for the cloud cover. The averaged zonal mean plot shows that the underestimation is well distributed over all latitudes with a minimum over the northern tropics.

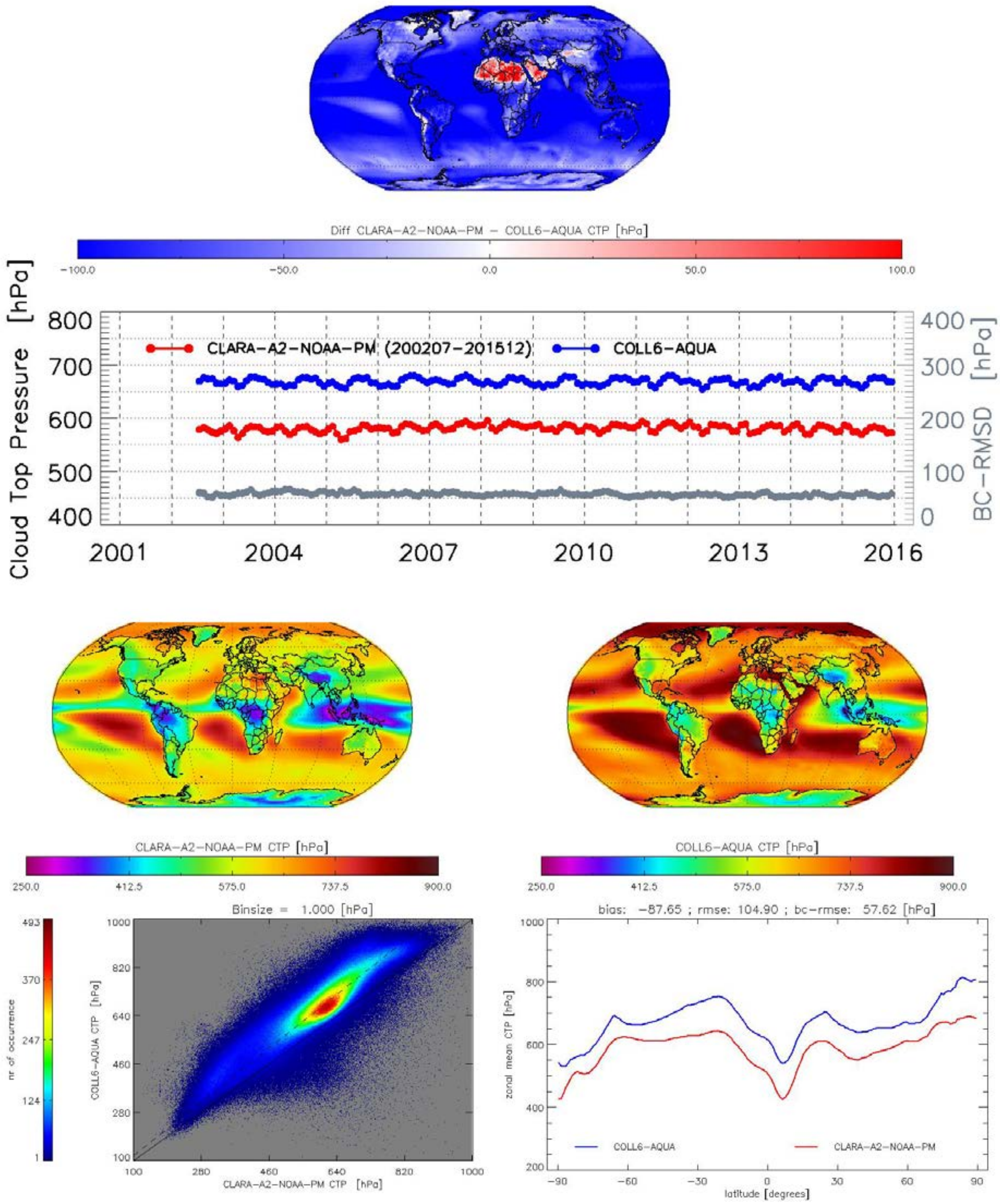


Figure 6.32 *Cloud top pressure comparison of CLARA-A2 afternoon satellites and MODIS collection 6 AQUA monthly means for the entire available time series 2002-2015. The top panel shows the difference plot, the panel below the time series and the bc-RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA-A2 left and MODIS right). The bottom left panel shows the 2D histogram of all datapoints in time and space and the bottom right panel the averaged zonal mean for CLARA-A2 in red and MODIS in blue.*

In Table 6.12 the bias and bc-RMSE is shown for the MODIS comparison. Here, also the comparison against MODIS TERRA is considered, showing comparable results.

Summary of results:

- Good agreement in overall vertical distribution of clouds and their geographical distribution
- CM SAF CTP values are generally lower (about -85 hPa), especially over ocean.
- Some positive deviations are found over the desert areas
- In the midlatitudes the deviation is generally small

Table 6.12 *Compliance matrix of found global CTP monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against MODIS results (consistency check).*

	CTP product requirements level-3 (MM)			MODIS/Aqua (2002-2015)	MODIS/Terra (2000-2015)
	Threshold	Target	Optimal		
Bias	80 hPa	50 hPa	30 hPa	-88 hPa	-84 hPa
bc-rms	120 hPa	100 hPa	80 hPa	58 hPa	60 hPa

6.2.3 Evaluation against ISCCP

6.2.3.1 Evaluation of CLARA-A2 CFC level-3 products

In this section CFC level-3 (monthly means) of CLARA-A2 is compared to ISCCP equivalents. The comparison is based on the full available time series. For ISCCP this is 1994-2008. Results are shown in Figure 6.33 using the same visualisation as previously in Figure 6.31. The top panel, illustrating the spatial distribution of the differences, shows an overestimation of CLARA-A2 CFC in the tropics and over high latitude oceans. However, the combined use of geostationary and polar orbiting satellites shows some discontinuities in the difference plot, as seen east of Africa and in the Southern Ocean. This makes it difficult to draw final conclusions.

The time series plot including the bc-RMSE show again a very smooth curve with a small seasonal cycle. After 2001 CLARA-A2 show a more pronounced negative bias in the time series. This coincides with the start of NOAA16. From this point onwards, in the afternoon orbit more than one satellite is operating. The averaged global maps show, that the global patterns match well. The 2D histogram shows a larger scatter than other comparisons but the zonal means of both data records match well again. Especially the region near the equator shows good agreement in the zonal mean. Outside of the tropics deviations increases. These are generally negative (smaller CFC for CLARA-A2) but an exception is seen over the southern mid-latitudes. The largest (negative) difference is seen over the Polar Regions.

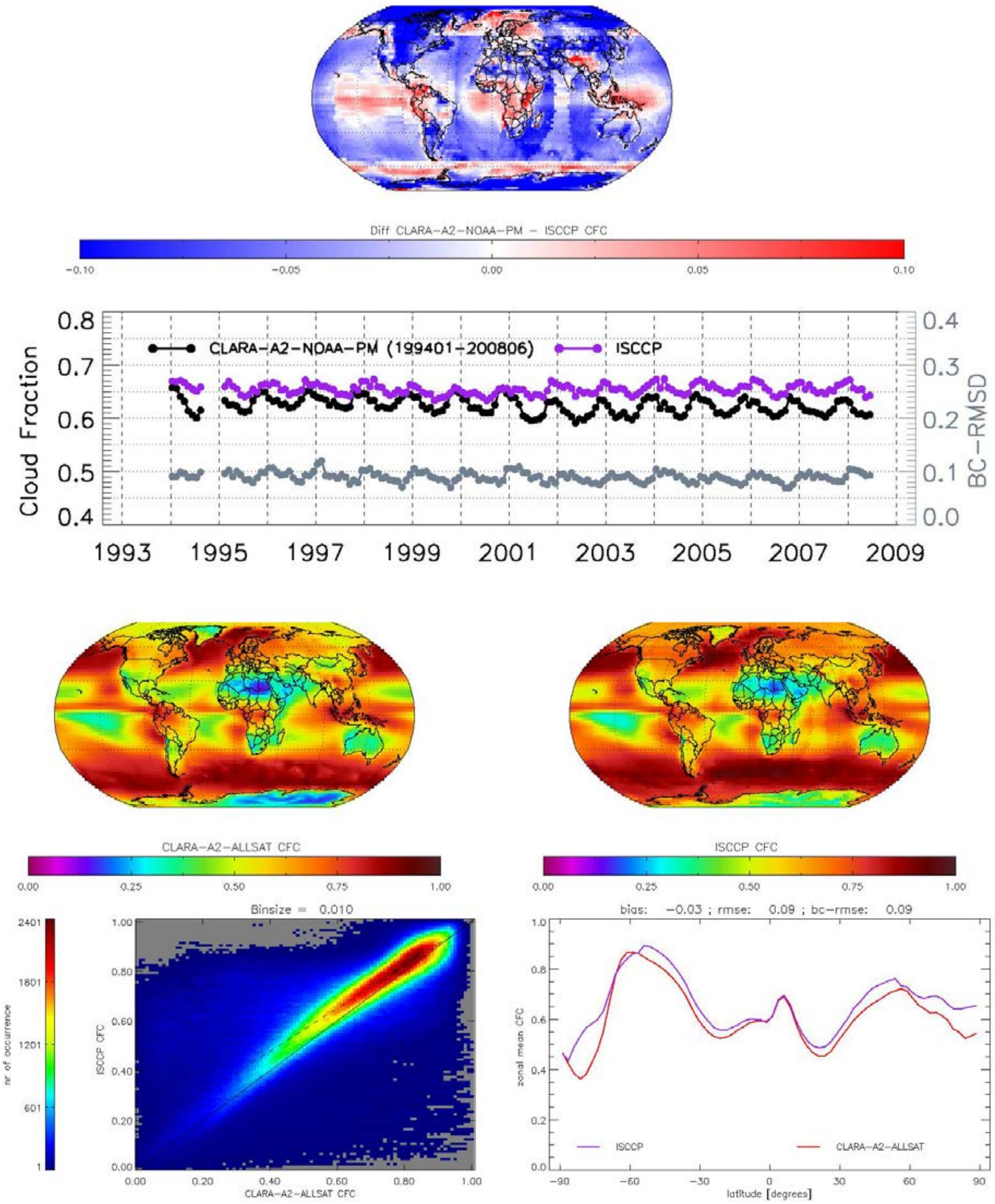


Figure 6.33 *Cloud cover comparison of CLARA-A2 and ISCCP monthly means for the entire available time series 1994-2009. The top panel shows the difference plot, the panel below the time series and the bc-RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA-A2 left and ISCCP right). The bottom left panel shows the 2D histogram of all datapoints in time and space and the bottom right panel the averaged zonal mean for CLARA-A2 in red and ISCCP in violet.*

Summary of results:

- Good agreement in the description of general global cloud features compared to ISCCP
- Overall lower CM SAF values with largest negative deviations seen for the North American continent and the Polar Regions.
- Positive deviations against ISCCP are found in the tropical ocean, as well as in the northern and southern sub-polar ocean.
- Global values for mean deviation and bias-corrected RMS are generally fulfilling target requirements

Table 6.13 *Compliance matrix of found global CFC monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against ISCCP observations (consistency check).*

	CFC product requirements level-3 (MM)			ISCCP (1994-2008)
	Threshold	Target	Optimal	
Bias	10 %	5 %	2 %	-3.4 %
bc-RMS	20 %	10 %	5 %	8.8 %

6.2.3.2 Evaluation of CLARA-A2 CTP level-3 products

In this section CTP level-3 (monthly means) of CLARA-A2 is compared to ISCCP equivalents. The comparison is based on the full available time series. For ISCCP this is 1994-2008. Results are shown in Figure 6.34 for the comparison of CLARA-A2 against ISCCP.

From top to bottom the global difference map, CLARA A2 – ISCCP, the time series, the averaged global plots, the 2D histogram and the averaged zonal mean is plotted in Figure 6.34. The difference map identifies mainly an overestimation of the cloud top pressure over land. Exceptions are only found over the Amazonas and the Saharan desert but also over the west part of subtropical Indian and Atlantic oceans and the south east subtropical pacific an overestimation can be observed. Tropical sea areas show an underestimation of the cloud top pressure. The stratocumulus areas off the west coast of Africa and America also show a clear underestimation by CLARA-A2.

The time series show a good agreement until 1999. After 1999 both data records deviate slightly, in which ISCCP decreases more than CLARA A2 increases. The change happens in 1999, which is the time NOAA15 becomes the prime morning satellite. The global maps of the averaged cloud top pressure reports a good agreement on the representation of the location and expansion of the different climate regions. But the magnitude of each data record deviates within a region. This is also indicated by the strong scatter in the histogram.

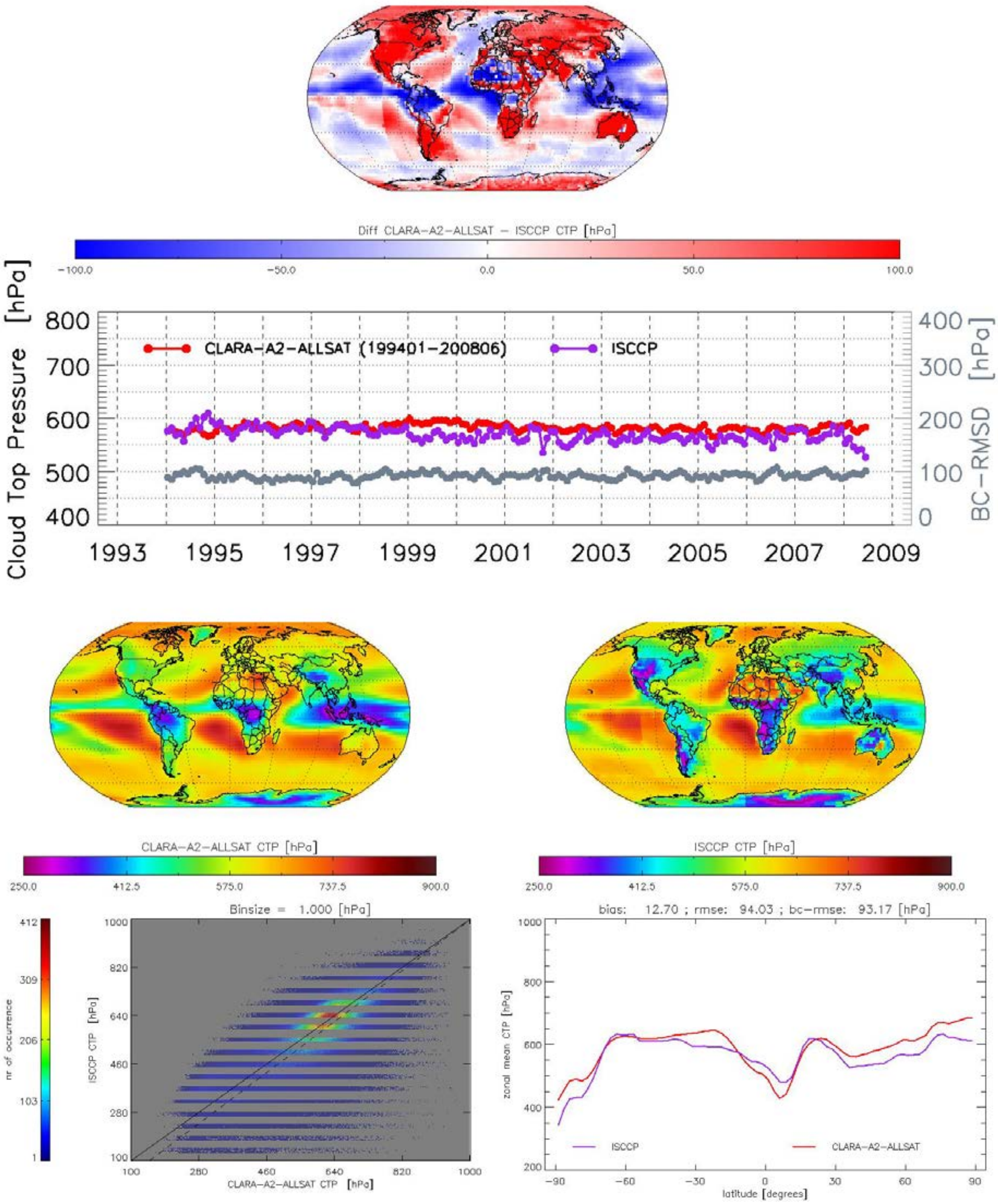


Figure 6.34 *This figure shows the cloud top pressure comparison of CLARA A2 and ISCCP monthly means for the entire available time series 1994-2009. The top panel shows the difference plot, the panel below the time series and the bc-RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA A2 left and ISCCP right). The bottom left panel shows the 2D histogram of all datapoints in time and space and the bottom right panel the averaged zonal mean for CLARA A2 in red and ISCCP in violet.*

In Table 6.14 the bias and bc-RMSE is shown for the ISCCP comparison. Both are within the target requirements.

Summary of results:

- CLARA-A2 shows a low bias against ISCCP compared to other data records, but also a high bc-RMSE
- Positive deviations are seen over most land surfaces
- Negative deviations are seen over deserts and the Amazonas
- Positive deviations are found over the tropical and subtropical oceans mainly towards the west coasts, but negative deviations prevail at the west coasts of Africa and America
- For the last 10 years in the time series, ISCCP values are generally up to 50 hPa lower than CLARA-A2
- Target requirements are generally fulfilled

Table 6.14 *Compliance matrix of found global CTP monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against ISCCP results (consistency check).*

	CTP product requirements level-3 (MM)			ISCCP (1993-2008)
	Threshold	Target	Optimal	
bias	80 hPa	50 hPa	30 hPa	13 hPa
bc-rms	120 hPa	100 hPa	80 hPa	93 hPa

6.2.4 Evaluation against PATMOS-x

6.2.4.1 Evaluation of CLARA-A2 CFC level-3 products

In this section CFC level-3 (monthly means) of CLARA-A2 is compared to PATMOS-x equivalents. The comparison is based on the full available time series. For PATMOS-x this is 1982-2009. Because PATMOS-x is only available for the prime satellites, in this section only these time ranges are considered for each satellite in the CLARA-A2 data record. Further, this comparison only considers afternoon satellites, because this enables to represent the full time series with a constant variability. This means all data are based on the same orbit and the same local observation time. The results for the analysis of the CLARA-A2 against PATMOS-x are shown in Figure 6.35.

The top panel, showing the difference plot, indicates generally lower CLARA-A2 cloud cover over ocean but higher cloud cover over most areas over land. Only over the Amazonas, in central Africa and Antarctica lower CFC over land is found. Over ocean the negative difference is strongest over the subtropical ocean and in the Arctic. All other sea areas show a small negative difference. The inner tropical belt shows a very low bias and a CFC that is nearly balanced in average. The second panel shows the time series. Here, CLARA-A2 shows as well a very homogenous time series without any obvious jumps. This is also expressed by the constantly low bc-RMSE. However, the PATMOS-x time series show an increase in CFC during the lifetime of NOAA-11 and NOAA-14 which is not yet fully understood.

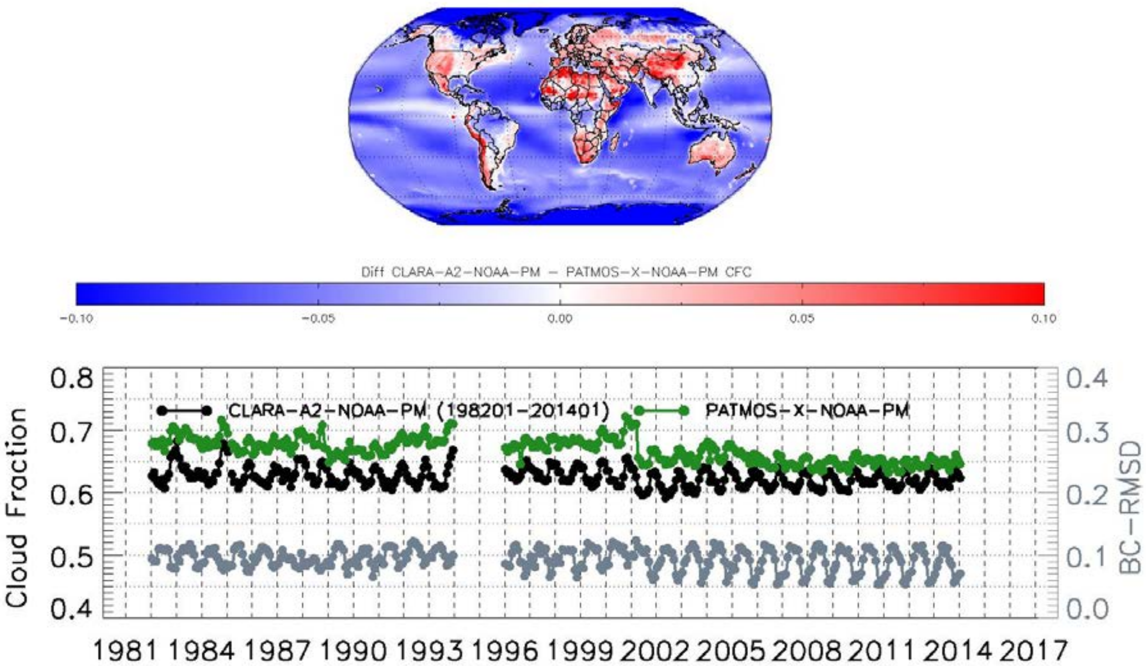
Interestingly, PATMOS-x and CLARA-A2 show diverging results near the Poles. But in this sense PATMOS-x seems to be an outlier if comparing also with MODIS and ISCCP results (see previous sub-sections).

Summary of results:

- Low bias for the global mean against PATMOS-x, but regionally larger deviations that balance each other negative deviations over the Ocean and positive deviation over land.
- Larger negative bias over subtropical ocean and nearly balanced results in the inner tropics
- Positive bias over land around the Mediterranean sea, northern China, and the Andes
- Very low deviation over the eastern continental areas (north-eastern America, eastern Brazil, Argentine, south-eastern China).

Table 6.15 *Compliance matrix of found global CFC monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against ISCCP observations (consistency check).*

	CFC product requirements level-3 (MM)			PATMOS-x (1994-2008)
	Threshold	Target	Optimal	
Bias	10 %	5 %	2 %	-4 %
bc-RMS	20 %	10 %	5 %	10 %



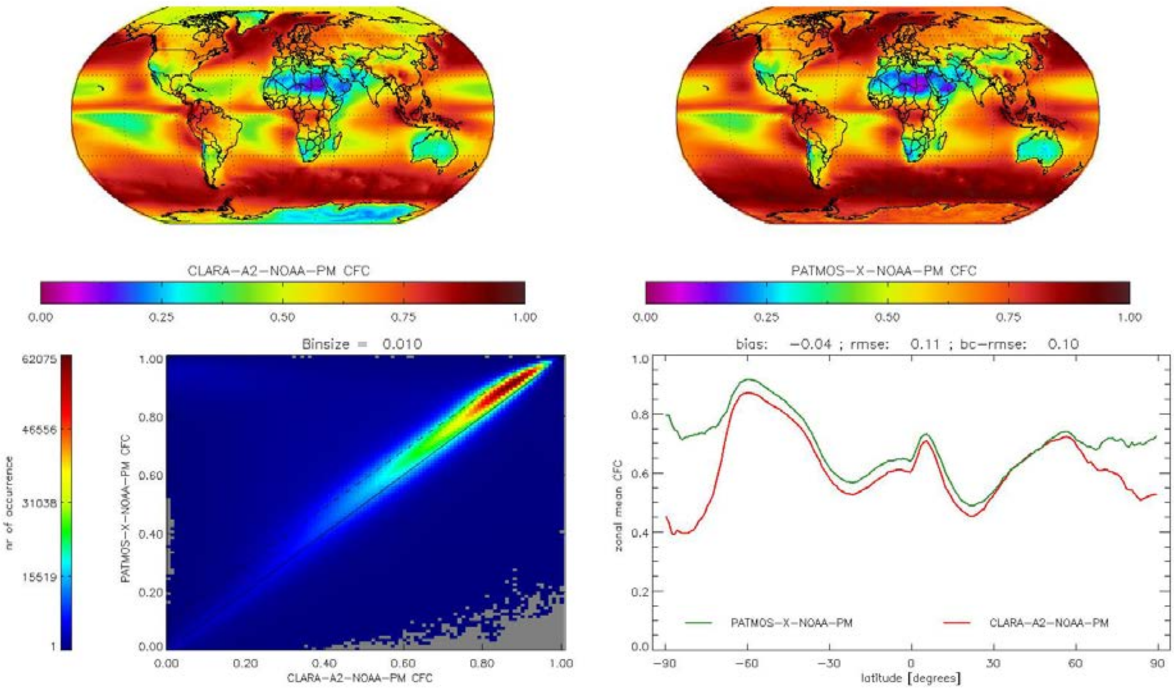


Figure 6.35 *Cloud cover comparison of CLARA-A2 afternoon prime satellites and PATMOS-x monthly means for the entire available time series 1982-2014. The top panel shows the difference plot, the panel below the time series and the bc- RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA-A2 left and PATMOS-x right). The bottom left panel shows the 2D histogram of all data points in time and space and the bottom right panel the averaged zonal mean for CLARA-A2 in red and PATMOS-x in green.*

6.2.4.2 Evaluation of CLARA-A2 CTP level-3 products

In this section CTP level-3 (monthly means) of CLARA-A2 is compared to PATMOS-x equivalents. The comparison is based on the full available time series. For PATMOS-x this is 1982-2015. Because PATMOS-x is only available for the prime satellites in this section the concentrated on these time ranges for each satellite. Further, this comparison only considers afternoon satellites, because this enables to represent the full time series with a constant variability. This means all data are based on the same orbit and, thus, has the same local observation time. The results for the analysis of the CLARA-A2 against PATMOS-x are shown in Figure 6.36.

The difference map in Figure 6.36 shows a lower cloud top pressure of CLARA-A2 over ocean, with slightly larger differences over the tropical and parts of the subtropical ocean, especially the stratocumulus regionsOver nearly the entire polar areas it a positive bias with lower clouds in CLARA A2 is shown.Over extra-tropical land a general positive deviation is found with higher clouds in PATMOS-x.

The time series match very well over the entire time range and globally only a very small bias is found. This is also specified in Table 6.16 by the bias and bc-RMSE. Here, the bias is very low but the bc-RMSE is relatively high.

Summary of results:

- In regions of very high clouds CLARA-A2 sees the clouds higher
- Overall, a very low bias
- Positive deviation over land and negative over ocean
- Larger deviation over marine stratocumulus regions
- Requirements are generally fulfilled

Table 6.16 *Compliance matrix of found global CTP monthly mean product characteristics with respect to the defined product requirements for accuracy and precision. Comparisons were made against PATMOS-x results (consistency check).*

	CTP product requirements level-3 (MM)			PATMOSx (1982-2009)
	Threshold	Target	Optimal	
bias	80 hPa	50 hPa	30 hPa	-18 hPa
bc-rms	120 hPa	100 hPa	80 hPa	85 hPa

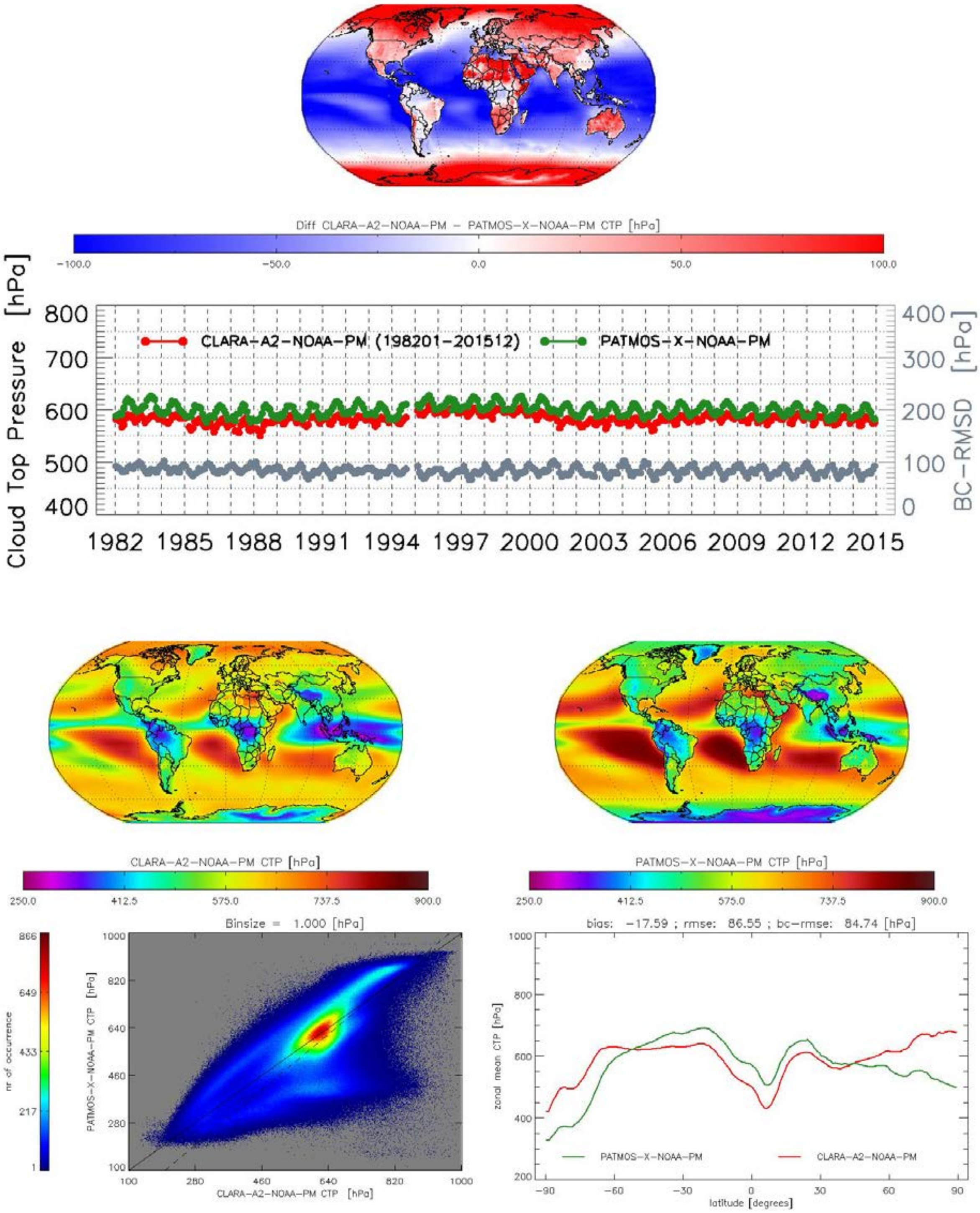


Figure 6.36 *Cloud top pressure comparison of results for CLARA-A2 afternoon satellites and PATMOS-x monthly means for the entire available time series 1982-2014. The top panel shows the difference plot, the panel below the time series and the bias-corrected RMSE. In the bottom quad panel the averaged global maps are shown in the top (CLARA-A2 left and PATMOS-x right). The bottom left panel shows the 2D histogram of all data points in time*

and space and the bottom right panel the averaged zonal mean for CLARA-A2 in red and PATMOS-x in green.

6.2.5 Evaluation of CPP products (CPH, LWP, IWP)

The evaluation of the microphysical cloud products, formally denoted Cloud Physical Products (CPP) and consisting of products CPH, LWP, IWP and additional products COT and REFF, has been performed using a common approach. This methodology is described here, before the individual product evaluation sub-sections.

The CPP level-3 products are compared with three data records: PATMOS-x, MODIS and ISCCP (see Sections 5.3, 5.4 and 5.5). In addition, LWP was validated with the O'Dell et al. (2008) climatology based on passive microwave observations (denoted UWisc, see Section 5.3). Table 6.17 shows these data records, their versions and the underlying instruments that were used.

Table 6.17 *Data records, their version and instruments that were used for the evaluation of the CPP products.*

Data record	version	Instruments
PATMOS-x	V05r03	NOAA-xx AVHRR
MODIS	Collection 6	Terra, Aqua
ISCCP	D1	Various GEO+LEO
UWisc	V4	SSM/I, TMI, AMSR-E

Evaluation of the CPP data records was done in terms of the bias and bias-corrected root mean square error (bc-RMS) compared to the other data records. In practice, the bias and bc-RMS for each month in the time series were calculated as the spatial mean and bias-corrected root-mean-square difference between two data records, respectively. In order to examine possible variations in the evaluation performance depending on latitude, these computations were performed for three different regions: globally, the tropics (30°S-30°N) and the areas excluding the tropics (90°S-30°S and 30°N-90°N). Apart from the overall and monthly bias and bc-RMS for each product, differences in spatial features between CLARA-A2 and the reference data records were also examined, based on pixel-level averages computed from the period when all data sets are available, and zonally averaged plots from the same period. This period comprises 2003-2007, with the first years being determined by the availability of a full year of MODIS data, and 2007 being the last year of ISCCP data available. Furthermore, time series plots of monthly values from the entire CLARA-A2 period were used, to give an overview of the level of agreement between the data sets. It should be noted that in all previous cases, CLARA-A2 level-3 data, which come at 0.25° × 0.25° spatial resolution, were first upscaled by simple averaging to 1° × 1°, in order to coincide with the PATMOS-x, MODIS and ISCCP spatial resolution. Then, results were computed using only pixels where all data sets had valid values.

It is also important to note that the CPP retrievals are based on different channel combinations for the different satellites. The non-absorbing channel used is always channel 1 (at 0.6 μm). The absorbing channel active during daytime is channel 3a (at 1.6 μm) on morning NOAA satellites and channel 3b (at 3.7 μm) on afternoon satellites, with the exception of NOAA-16 for the period 01/2001-04/2003, when channel 3a was active. Due to these differences, the CPP product evaluation is split into morning and afternoon satellites. For the former, NOAA-17, MetOp-A and MetOp-B are considered; for the latter, NOAA-7, -9, -11, -14, -16, -18 and -19 are considered. Thus, differences between CLARA-A2 products for morning and afternoon may reflect the differences in channel combination rather than those related to the time of day. For months when data from more than one satellite are available, which is quite common after 2000, average values were used, for purposes of consistency with PATMOS-x, where no satellite discrimination is performed in level-3 data, and ISCCP, where multiple satellites are used for the creation of the data set.

Since the CPP retrieval algorithm is restricted to solar zenith angles within 84°, and the level-3 processing takes into account data from solar zenith angles up to 75°, results obtained by the twilight satellites (NOAA-12 and NOAA-15, with local overpass times between 5:00 AM/PM and 7:30 AM/PM) have reduced spatial coverage. Furthermore, no reference products are available in the same time intervals, with the exception of cloud phase from PATMOS-x. Due to these limitations, evaluation of CPP products from these two satellites was limited to day-only cloud phase against PATMOS-x.

6.2.5.1 Cloud phase evaluation against PATMOS-x, MODIS, and ISCCP

Cloud phase is expressed here as the fraction of liquid water cloud amount over the total cloud amount. Two cloud phase data records are available in CLARA-A2, one derived from both day and night measurements (CPH) and the other from day-only measurements (CPH_Day) of each satellite. For PATMOS-x and ISCCP, corresponding data sets were created by dividing the liquid water cloud amount with the total (liquid + ice) amount, while in the case of MODIS an additional mixed/undetermined phase is included in the total amount.

Figure 6.37 shows the averaged spatial distribution of CPH from CLARA-A2, PATMOS-x, MODIS and ISCCP, over the period when all data sets overlap, for afternoon satellites. Patterns in all cases are very similar, highlighting e.g. the marine stratocumulus cloud areas at the eastern margins of the subtropical oceans, where liquid cloud fraction reaches values close to one, and Polar Regions, namely Greenland and Antarctica, where it acquires minimum values. Overall, ISCCP exhibits more extended areas of low CPH values in higher latitudes, compared to the other data sets, which leads to the corresponding low values depicted in both the zonally averaged CPH (Figure 6.38) and the spatially averaged time series (Figures 6.39c and 6.40c). Spatially averaged CPH from morning satellites is not shown here, due to its similarity with the results presented in Figure 6.37.

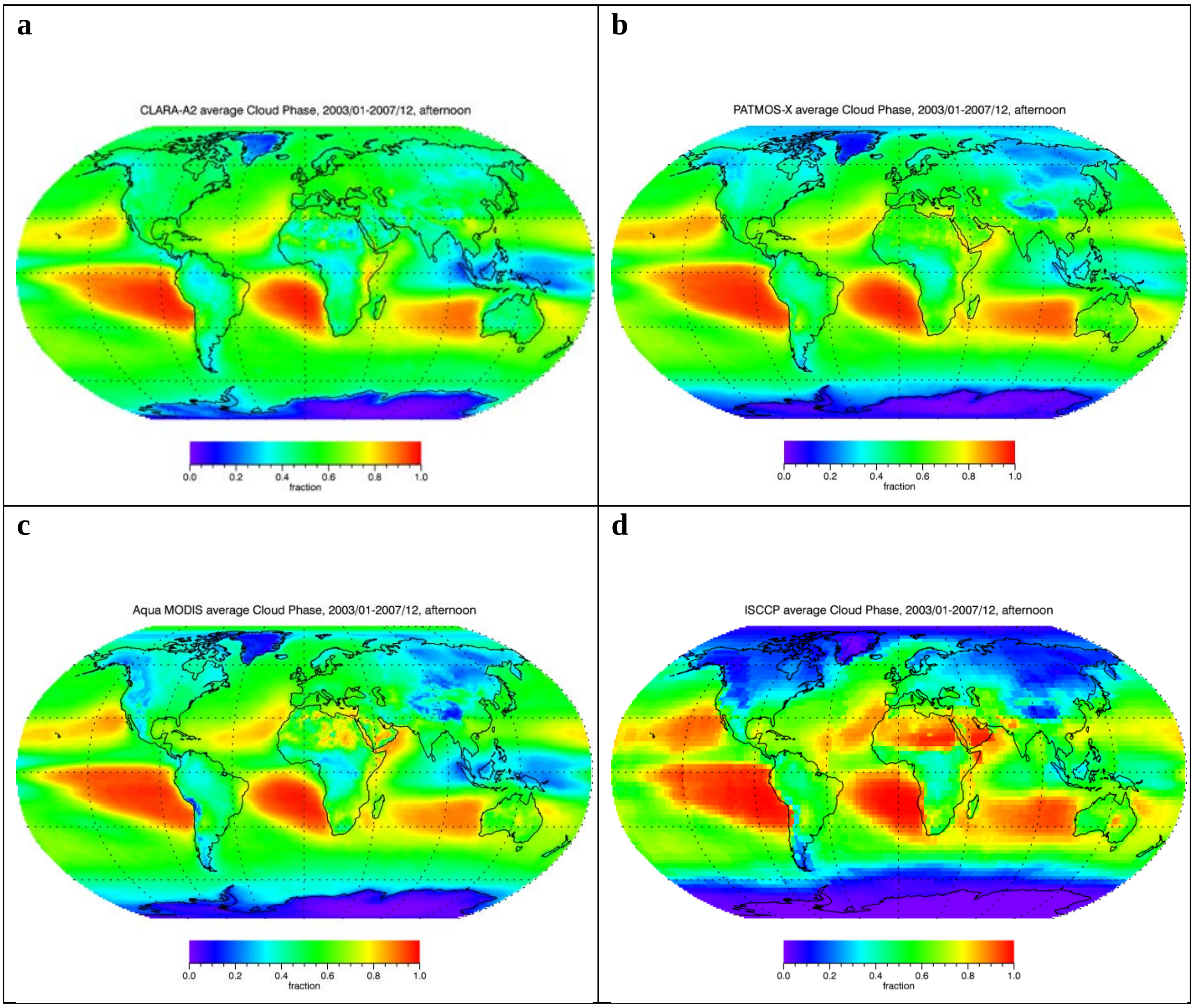


Figure 6.37 *Spatial distribution of afternoon CPH (expressed as liquid cloud fraction) from CLARA-A2 (a), PATMOS-x (b), Aqua MODIS (c) and ISCCP (d), averaged over the period when all data records were available (01/2003-12/2007).*

The zonally averaged CPH, presented in Figure 6.38 separately from morning and afternoon satellites, shows that CLARA-A2 is within the target accuracy (lighter shaded areas) with respect to PATMOS-x and MODIS in almost every latitude, while ISCCP tends to acquire higher CPH values in the tropics and lower in higher latitudes, as mentioned before. This is also apparent in the corresponding time series plots (Figures 6.39 and 6.40). The time series plots for morning satellites (Figure 6.39) also reveal the impact of the Terra MODIS band 29 (8.6 μm) radiometric calibration drift issue. As described in the MODIS Atmosphere products website (http://modis-atmos.gsfc.nasa.gov/validation_35.html), the root problem is a gradual warming in the Terra MODIS band 29 over the years that is not being captured by the on-board calibration systems and has not been corrected yet. The effect on monthly level-3 cloud phase data originates from the use of this channel in the level-2 cloud mask product (MOD35). It is, however, irrelevant to the other MODIS products used here, since they are based on the optical properties retrieval algorithm. Since this issue affects data after 2007, results presented in Figures 6.37, 6.38, 6.41 and 6.42 are not affected. For the same reason, the bias and bc-RMS analysis was based on Terra MODIS data until the end of 2007.

Apart from the previously mentioned issue, the globally averaged time series analysis of CPH (Figures 6.39 and 6.40) reveals a very good agreement between all data sets, with corresponding biases being within the target threshold, except for the ISCCP CPH estimated over the tropics (Figure 6.39b and parts of Figure 6.40b).

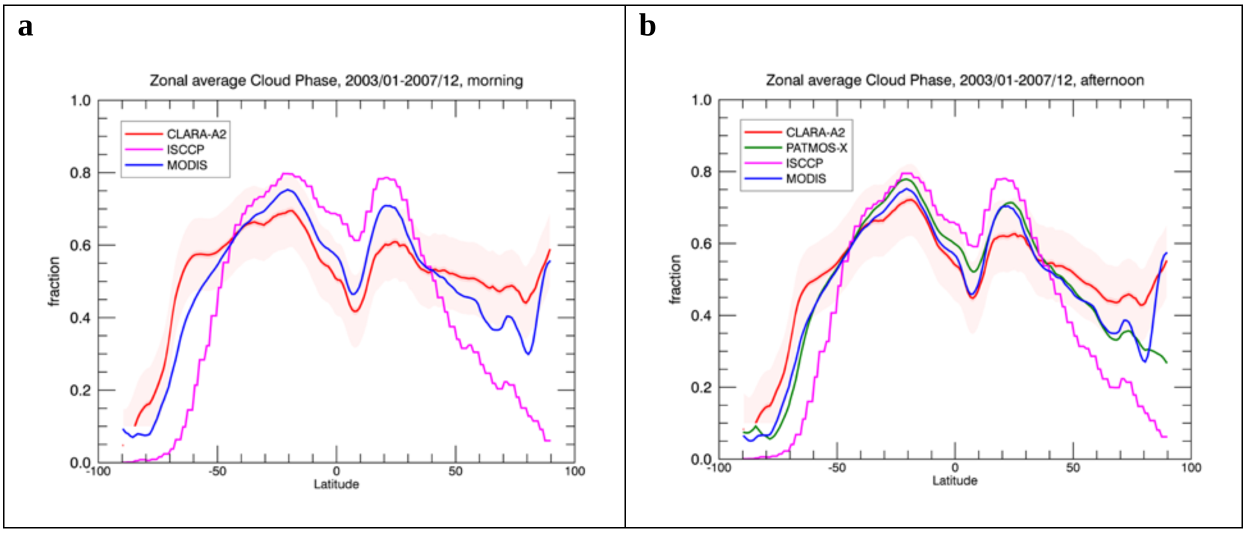


Figure 6.38 *Zonal average CPH for morning (a) and afternoon (b) satellites, for CLARA-A2, PATMOS-x, MODIS and ISCCP, computed from corresponding averages from their common periods (given in Figure 6.37 caption). The shaded area around CLARA-A2 curves denotes the target accuracy. Optimal accuracy is equal to 0.01.*

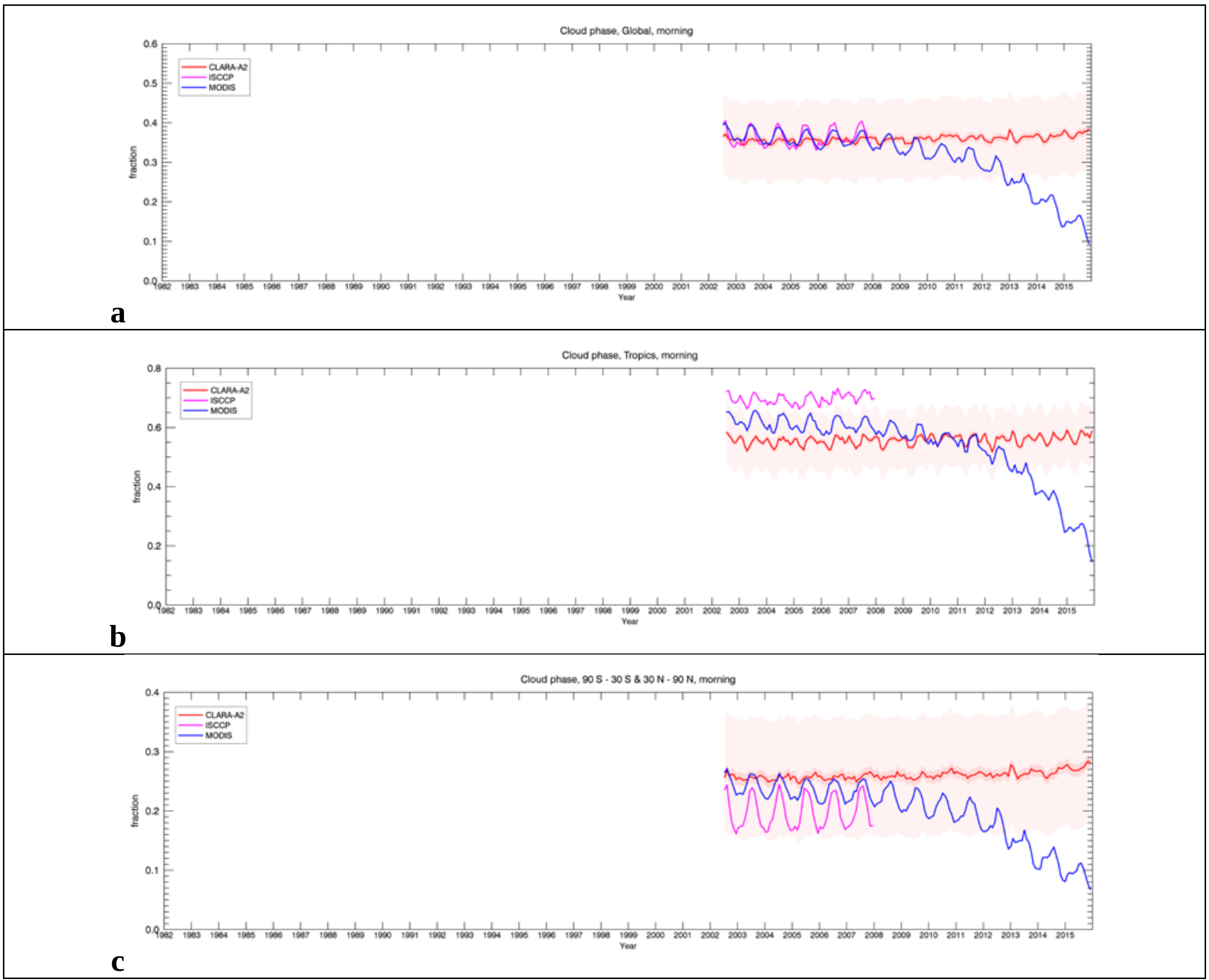


Figure 6.39 *Time series of the morning CPH from CLARA-A2, MODIS and ISCCP, averaged over the globe (a), the tropics (b) and the areas excluding the tropics (c). The darker and lighter shaded areas around the CLARA-A2 curves denote the optimal and target accuracies, respectively.*

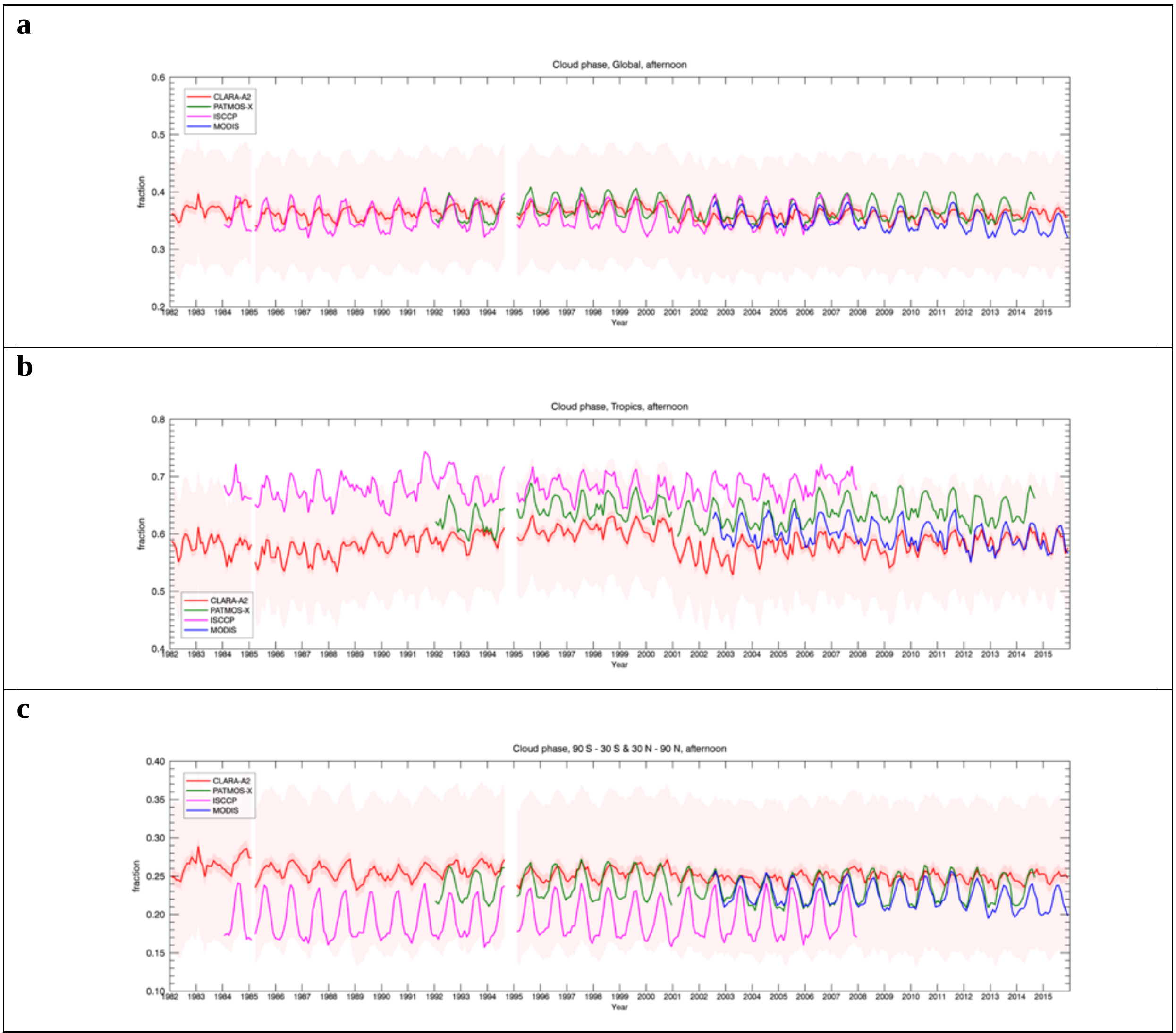


Figure 6.40 *As in Figure 6.39, for afternoon satellites.*

6.2.5.2 Evaluation of day-only CPH (CPH_Day)

Evaluation results of the CPH_Day are very similar to the CPH product in terms of their spatial characteristics (Figure 6.41). The main difference is the systematically lower CPH_Day by ISCCP afternoon data, compared to the other data sets, which is apparent in all latitudes (Figure 6.42b) and during the entire period covered (Figure 6.44).

Figure 6.42 shows that CLARA-A2 CPH_Day is within the target accuracy threshold (lighter shaded areas) with respect to PATMOS-x and MODIS in almost all latitudes, while ISCCP also lies within this threshold (morning, Figure 6.42a) or at its edge (afternoon, Figure 6.42b) in the zone between 50°S and 30°N. The time series analysis for morning satellites (Figure 6.43) reveals an agreement within the target threshold between CLARA-A2, ISCCP and MODIS until 2007, when the Terra MODIS calibration drift issue becomes apparent. In the

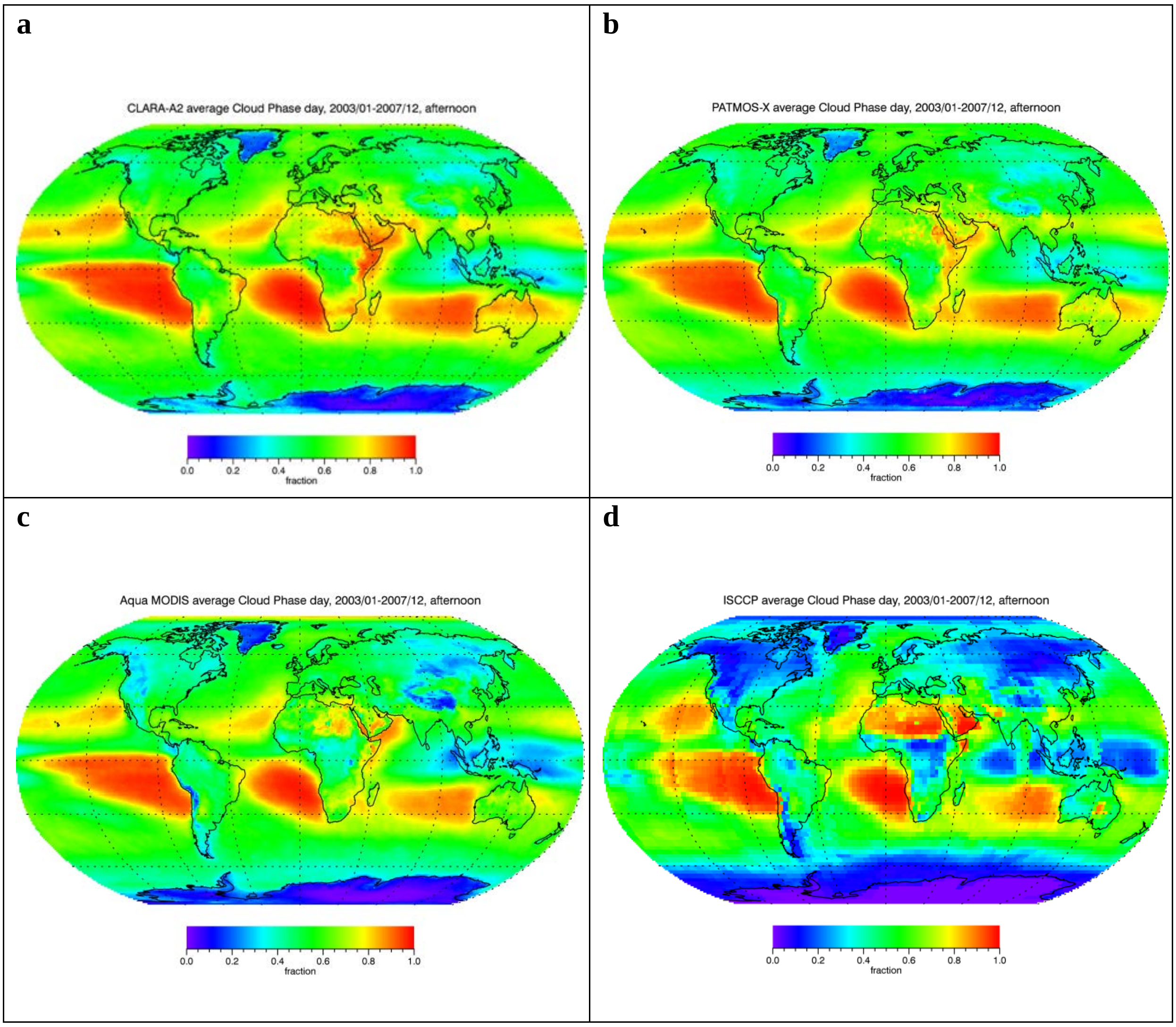


Figure 6.41 *As in Figure 6.40, for CPH_Day.*

afternoon CPH_Day case, some irregularities that appear in the CLARA-A2 time series (Figure 6.44), should probably be attributed to satellite transitions, e.g. the slightly higher values in the tropics during 1995-2001 (Figure 6.44b), which coincides with NOAA-14 operational period.

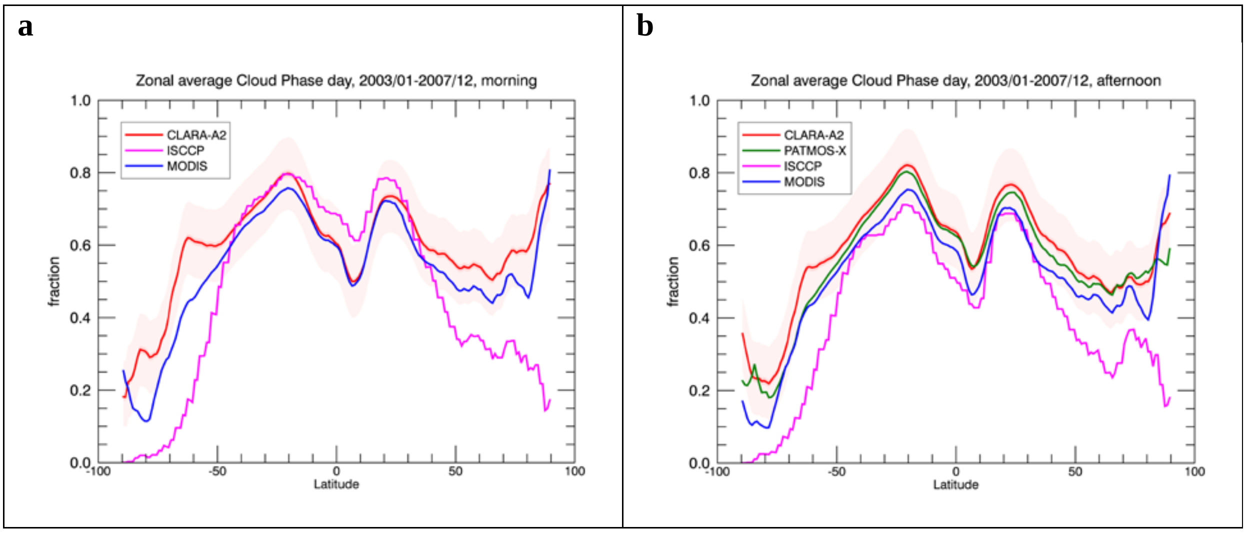


Figure 6.42 *As in Figure 6.38, for CPH_Day.*

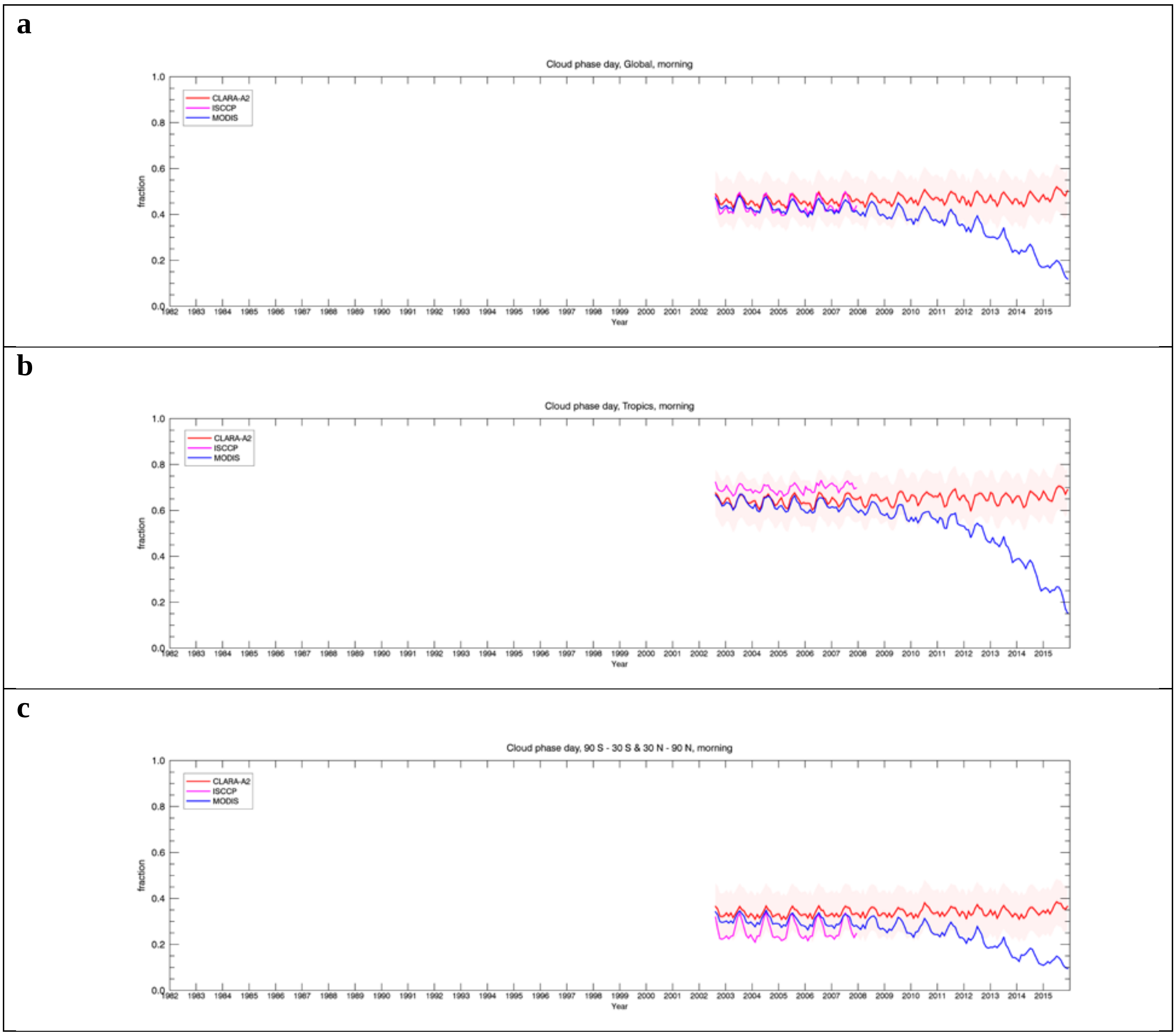


Figure 6.43 *As in Figure 6.39, for the morning CPH_Day.*

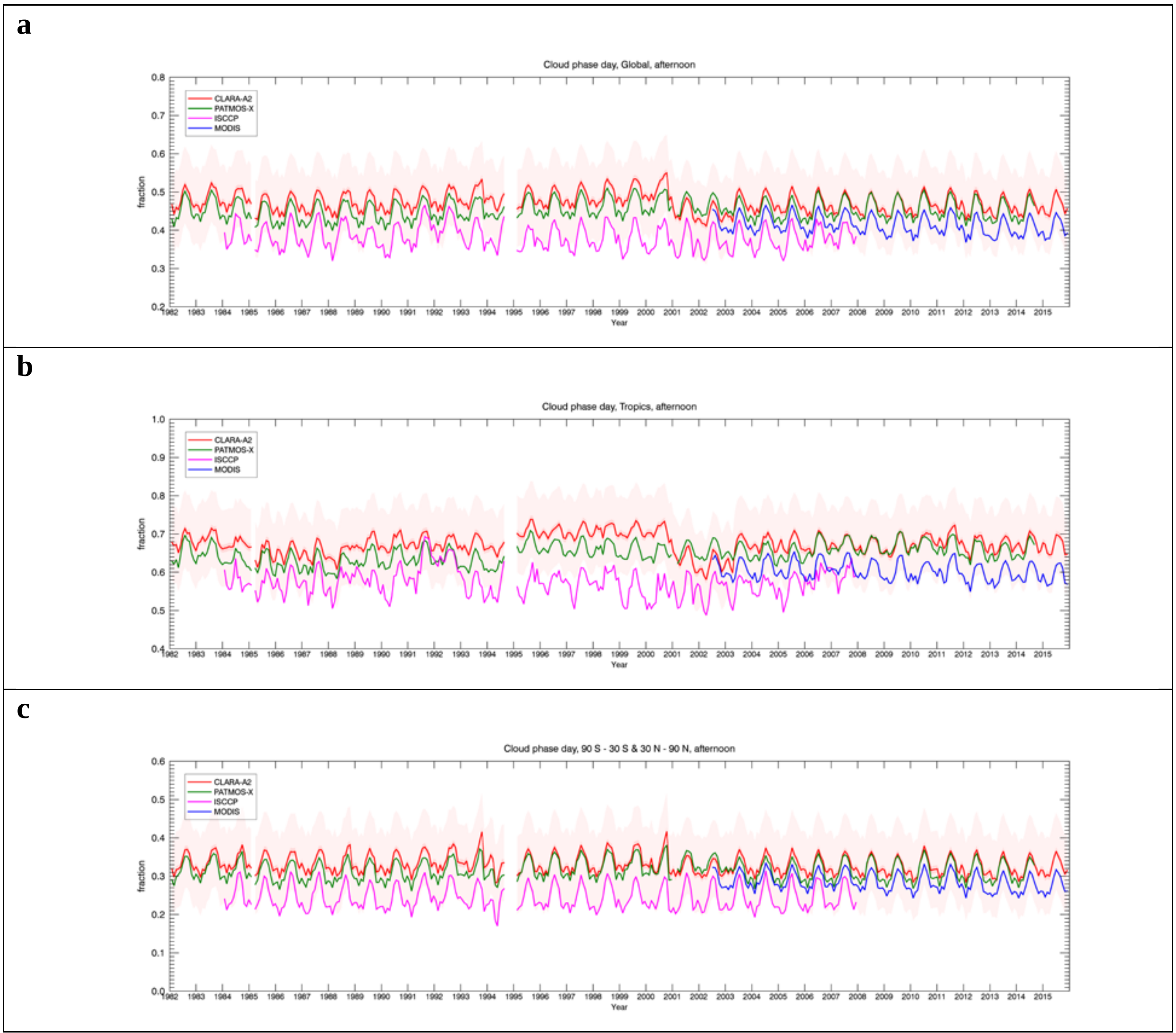


Figure 6.44 *As in Figure 6.43, for the afternoon satellites.*

6.2.5.3 Evaluation of CPH_Day in twilight conditions

As mentioned before, only CPH_Day data from PATMOS-x were available for the evaluation of twilight satellites (NOAA-12 and -15) products. Results are shown in Figure 6.45. While monthly differences between CLARA-A2 and PATMOS-x are within the target accuracy in all regions examined, and the average bias (0.01) and bc-RMS (0.11) are within the optimal (0.01) and target (0.2) accuracies, there are several irregularities in the time series, caused by reduced spatial coverage. As shown in Figure 6.45d, orbital drifts of both NOAA-12 and NOAA-15 cause large drops in the areas covered on a monthly basis. Since comparisons between CLARA-A2 and PATMOS-x are based on commonly covered areas, the data sets agree well even in these cases. These values, however, cannot be considered representative of the broader areas examined.

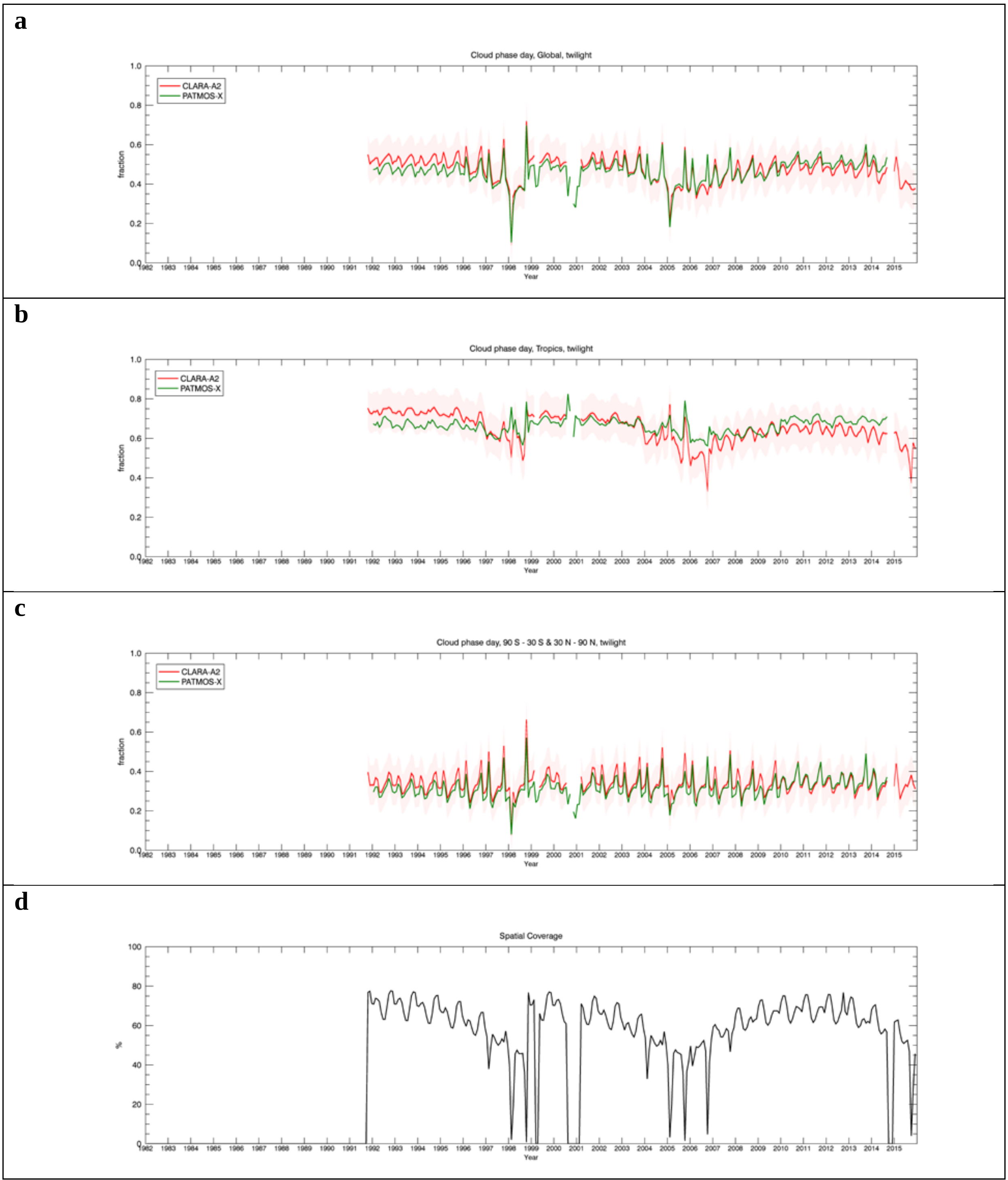


Figure 6.45 *Time series of the twilight CPH_Day from CLARA-A2 and PATMOS-x, averaged over the globe (a), the tropics (b) and the areas excluding the tropics (c). The lighter shaded areas around the CLARA-A2 curves denote the target accuracy, while optimal accuracy is 0.01. The spatial coverage (in %) of the averaged data in the global case is shown in (d).*

6.2.5.4 Summary of overall CPH validation results

Tables 6.18 and 6.19 summarize the evaluation of CLARA-A2 CPH and CPH_Day in terms of the globally averaged bias and bc-RMS with respect to PATMOS-x, MODIS and ISCCP, calculated from the full available time series of each data set, except for Terra CPH, where the 2003-2007 period was used, due to the degradation issue. Results are presented separately for morning and afternoon satellites, and compliance with the predefined requirements is indicated by simple YES or NO statements.

- Compared to PATMOS-x, where the evaluation included only afternoon satellite data, optimal requirements of bias and bc-RMS are fulfilled, except for the bias in CPH_Day, which is below the target threshold.
- Compared to MODIS, both the bias and bc-RMS are most of the times below the optimal threshold for both CPH and CPH_Day; in the case of the bias in CPH_Day, the target requirement is fulfilled.
- Compared to ISCCP, target requirements are always fulfilled. In the case of CPH bias, the optimal threshold is also achieved.

Table 6.18 *Overall requirement compliance of the CLARA-A2 CPH product with respect to the Mean Error and the bias-corrected RMS (bc-RMS). Consistency checks marked in blue.*

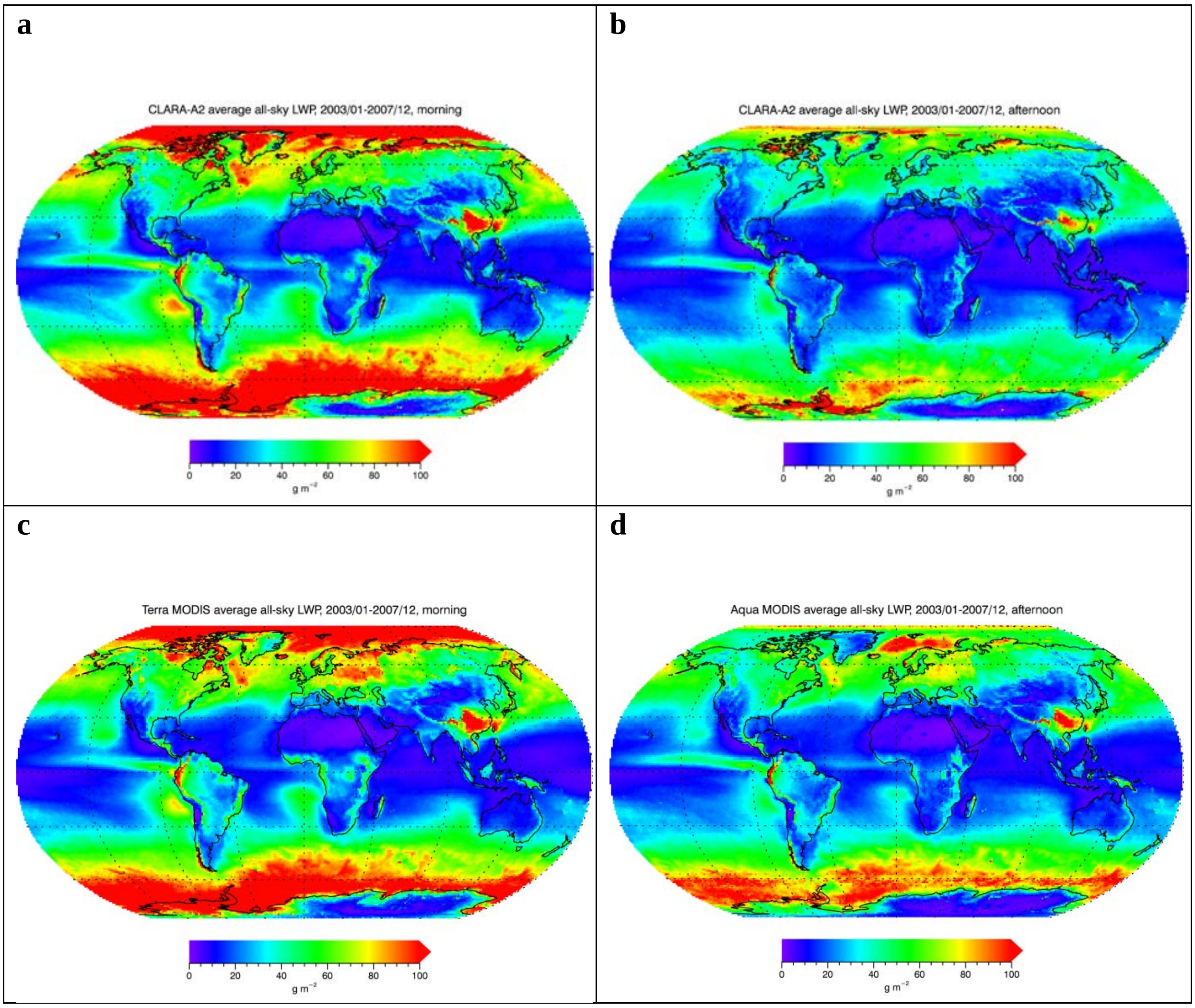
Reference data set	Mean Error (Morning/Afternoon)	Fulfilling Threshold requirements (0.2)	Fulfilling Target Requirements (0.1)	Fulfilling Optimal Requirements (0.01)
PATMOS-x	--/-0.006	--/YES	--/YES	--/YES
MODIS	-0.007/0.009	YES/YES	YES/YES	YES/YES
ISCCP	-0.009/0.008	YES/YES	YES/YES	YES/YES
	bc-RMS	Fulfilling Threshold requirements (0.4)	Fulfilling Target Requirements (0.2)	Fulfilling Optimal Requirements (0.1)
PATMOS-x	--/0.074	--/YES	--/YES	--/YES
MODIS	0.091/0.080	YES/YES	YES/YES	YES/YES
ISCCP	0.156/0.134	YES/YES	YES/YES	NO/NO

Table 6.19 *As in 6.18, for the CPH_Day product.*

Reference data set	Mean Error (Morning/Afternoon)	Fulfilling Threshold requirements (0.2)	Fulfilling Target Requirements (0.1)	Fulfilling Optimal Requirements (0.01)
PATMOS-x	--/0.020	--/YES	--/YES	--/NO
MODIS	0.027/0.052	YES/YES	YES/YES	NO/NO
ISCCP	0.023/0.094	YES/YES	YES/YES	NO/NO
	bc-RMS	Fulfilling Threshold requirements (0.4)	Fulfilling Target Requirements (0.2)	Fulfilling Optimal Requirements (0.1)
PATMOS-x	--/0.062	--/YES	--/YES	--/YES
MODIS	0.079/0.075	YES/YES	YES/YES	YES/YES
ISCCP	0.138/0.141	YES/YES	YES/YES	NO/NO

6.2.5.5 Liquid water path evaluation against PATMOS-x, MODIS, and ISCCP

Figure 6.46 shows the spatial distribution of all-sky LWP from CLARA-A2, MODIS, ISCCP and PATMOS-x, computed as averages from the full years when all data sets were available (2003-2007), separately for morning and afternoon satellites. Spatial features are similar in all data sets, with the lowest all-sky LWP values occurring in areas around the tropics, and the highest in Polar Regions. It is also worth noting that the feature of decreased LWP in the afternoon compared to morning is captured in all data sets. Notice however, that differences between morning and afternoon LWP can also to a considerable part be attributed to the different SWIR channels (1.6 µm and 3.7 µm, respectively) active on the corresponding AVHRR instruments (see also Section 6.2.5.6). CLARA-A2 agrees well with MODIS in almost all latitudes and in both morning and afternoon, while there are overall lower values in ISCCP and higher in the Arctic region by PATMOS-x. These results are verified by the latitudinal averages of all data sets, presented in Figure 6.47. The shaded areas around the CLARA-A2 curves, denoting the optimal (darker) and target (lighter) accuracies, reveal that all afternoon data sets lie within at least the target accuracy in an area extending the tropics, while morning CLARA-A2 and MODIS are in very good agreement in all latitudes.



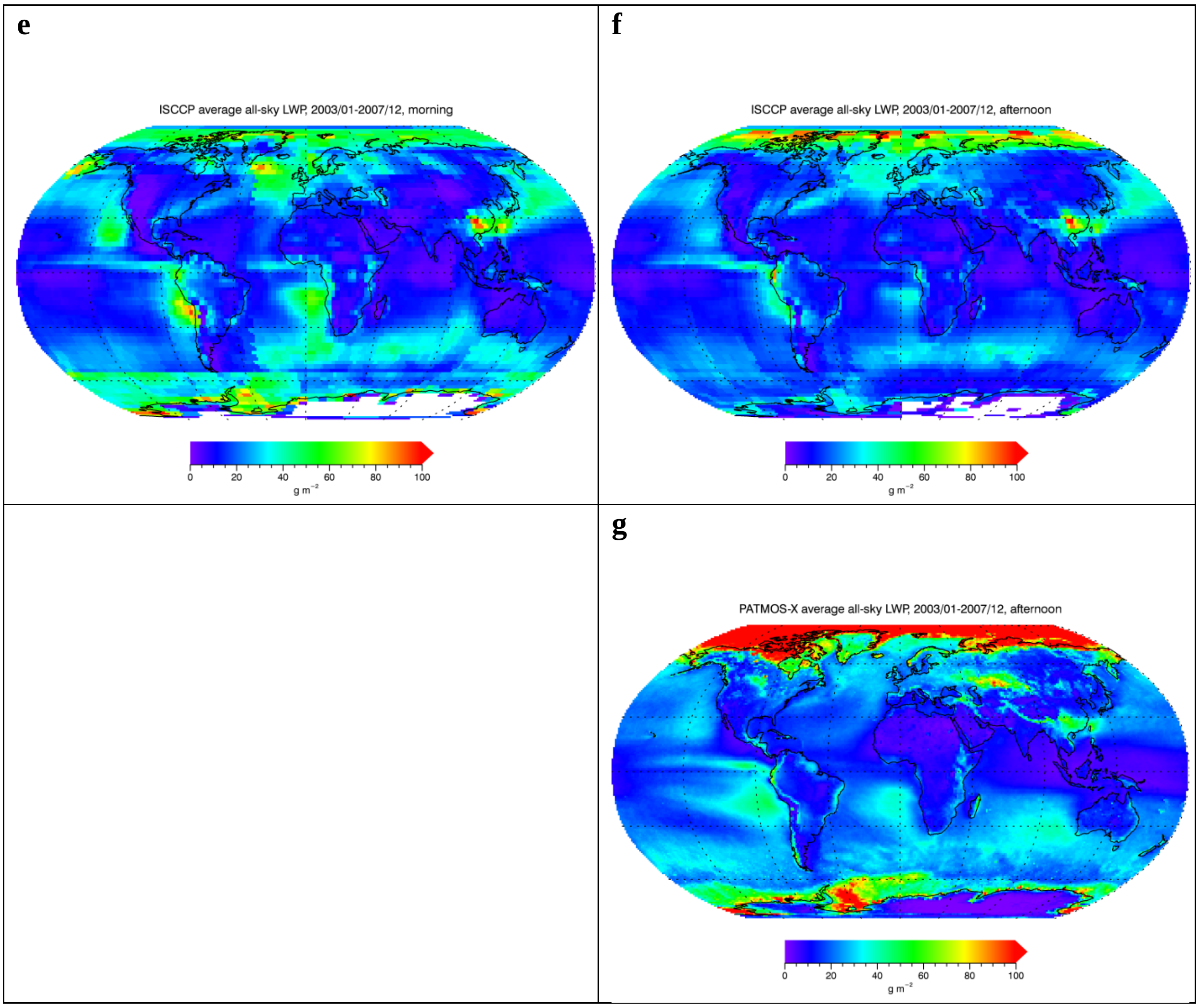


Figure 6.46 *Spatial distribution of the all-sky LWP from CLARA-A2 (a, b), MODIS (c, d), ISCCP (e, f) and PATMOS-x (g), separately for morning (left column) and afternoon (right column) satellites, averaged over the period when all data records were available (01/2003-12/2007). The all-sky LWP from PATMOS-x morning satellites was not available.*

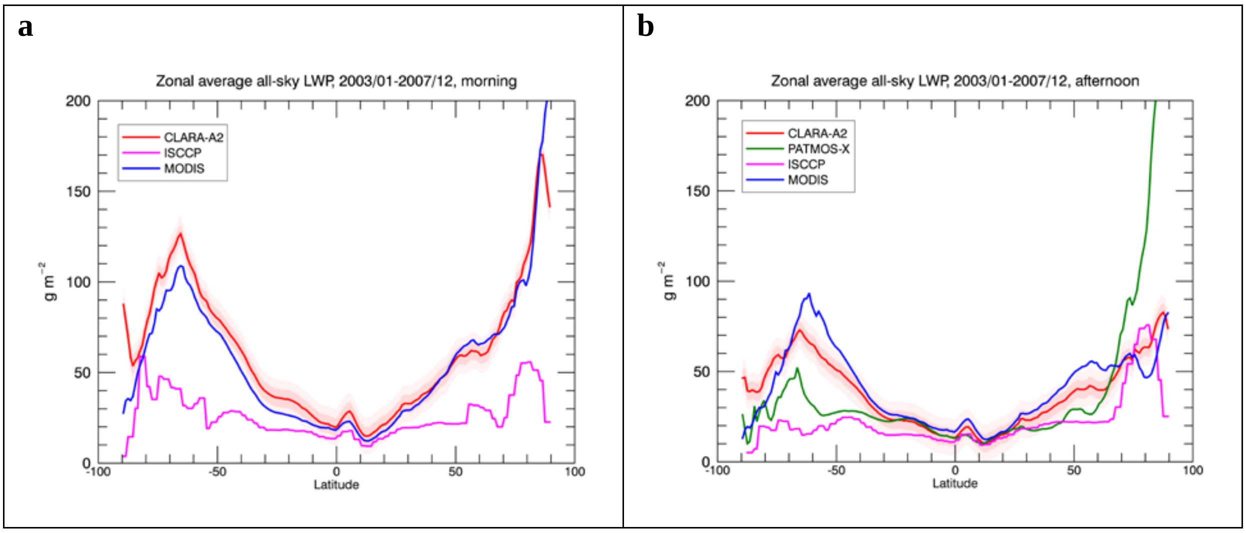


Figure 6.47 *Zonal average all-sky LWP for morning (a) and afternoon (b) satellites, for CLARA-A2, PATMOS-x, MODIS and ISCCP, computed from corresponding averages from their common period (01/2003-12/2007). The darker and lighter shaded areas around the CLARA-A2 curves denote the optimal and target accuracies, respectively.*

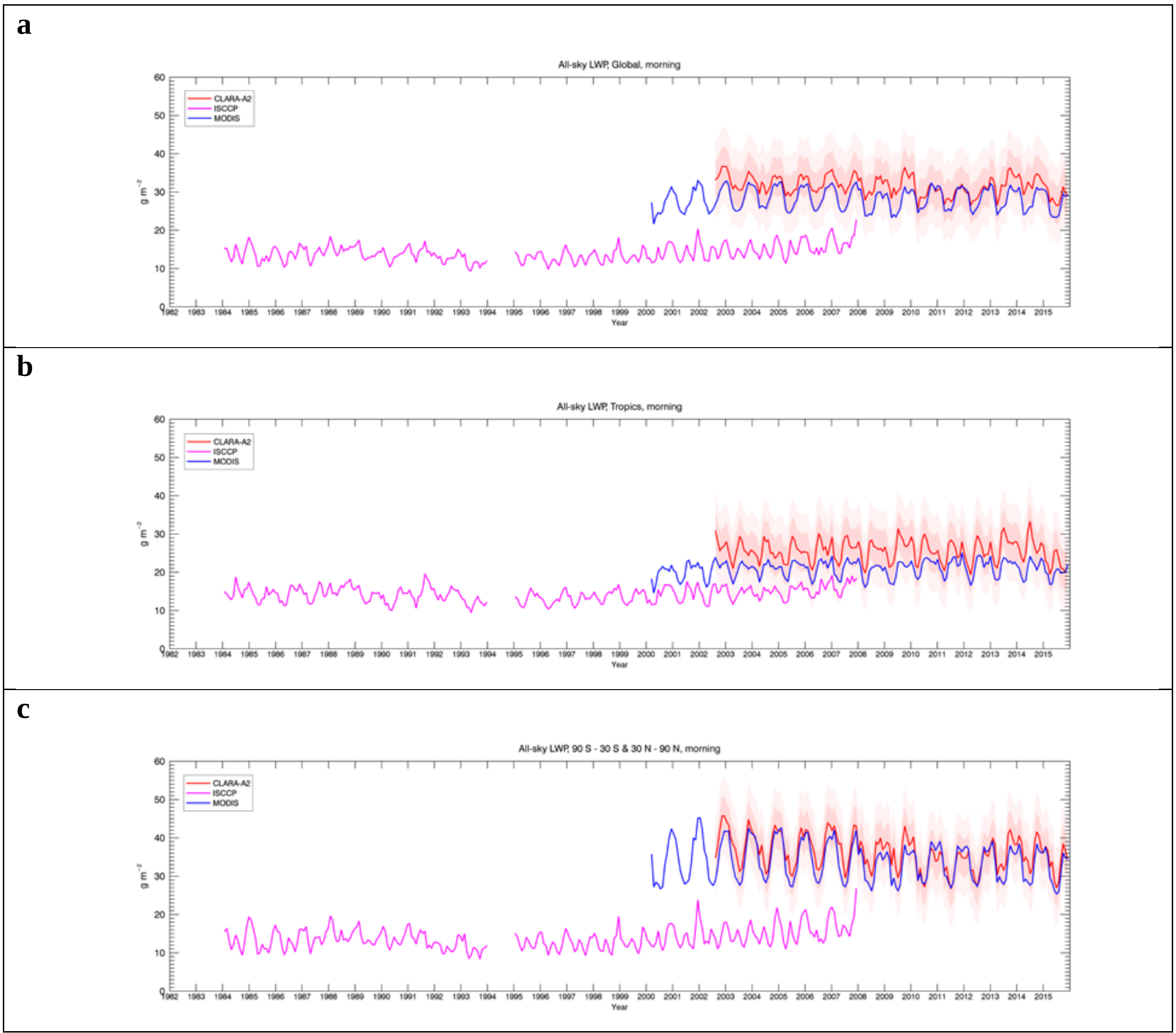
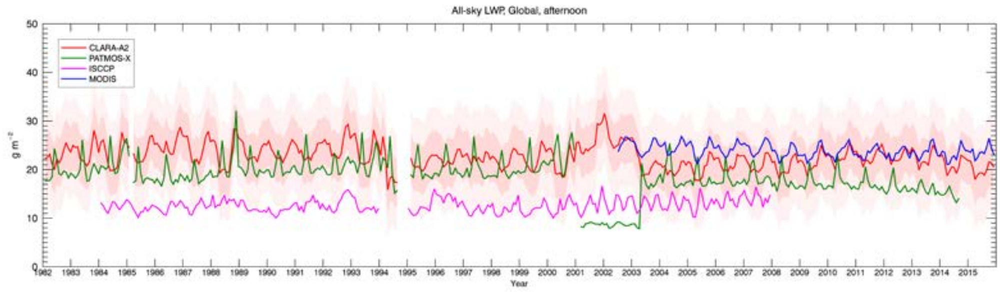


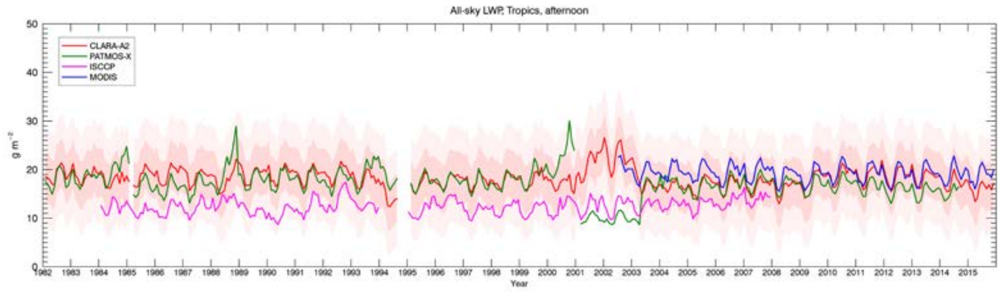
Figure 6.48 *Time series of the morning all-sky LWP from CLARA-A2, MODIS and ISCCP, averaged over the globe (a), the tropics (b) and the areas excluding the tropics (c). The darker and lighter shaded areas around the CLARA-A2 curves denote the optimal and target accuracies, respectively.*

The time series of the morning satellites (Figure 6.48) shows that CLARA-A2 is in very good agreement with Terra MODIS in all areas examined; their difference fluctuates around the optimal accuracy threshold and most of the times is lower, when tropical areas are not included (Figure 6.48c). ISCCP, however, lies below CLARA-A2 and MODIS, with the minimum differences occurring in the tropics.

a



b



c

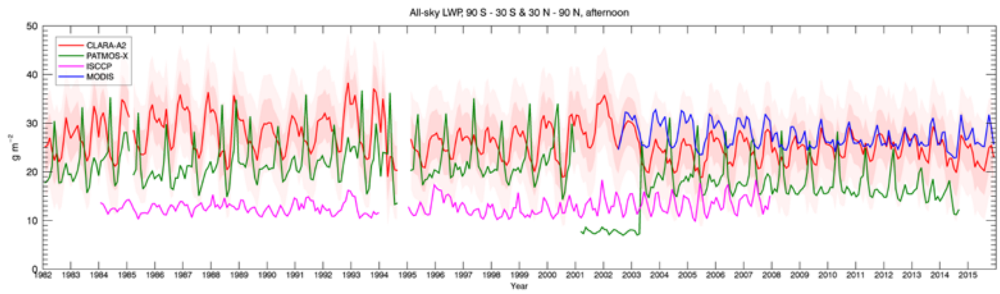


Figure 6.49 *As in Figure 6.48 but for the afternoon satellites.*

Results are similar in the case of afternoon satellites (Figure 6.49); the ISCCP all-sky LWP is systematically lower, with differences from CLARA-A2 within the 20 g m^{-2} threshold accuracy most of the times, especially during the last years of ISCCP availability. The best agreement occurs between CLARA-A2 and Aqua MODIS. CLARA-A2 also agrees well with PATMOS-x in the tropics (Figure 6.49b), while this agreement deteriorates as higher latitude regions are included (Figure 6.49a and c), where PATMOS-x all-sky LWP values are lower, with the exception of sharp peaks every spring and the higher values in the Arctic (Figure 6.49g). The latter characteristic, however, is not apparent in the time series due to the area weighted averaging. Another PATMOS-x characteristic worth noting is a small decline in all-sky LWP values occurring after 2012. A similar decline also appears in CLARA-A2 after 2014.

The irregularity in the CLARA-A2 time series, occurring during 2001-2003, should be attributed to the AVHRR channel 3a on board NOAA-16, which was on until April 2003. It is worth noting that, while channel 3a data led to higher all-sky LWP in CLARA-A2, compared to the rest of the period when channel 3b was on, the same channel switch causes a drop in all-sky LWP estimated by PATMOS-x, appearing from 2001 until early 2003.

6.2.5.6 Evaluation of liquid Cloud Optical Thickness (COT) and Effective Radius (REFF)

Further examination of the liquid COT and REFF, which are used for the computation of LWP, reveal that COT dominates in the LWP behavior. Figure 6.50 shows the time series of the globally averaged all-sky liquid COT and corresponding REFF from afternoon satellites. In the all-sky liquid COT case, similarities with the corresponding LWP (Figure 6.49a) are apparent, including the good agreement of CLARA-A2 with MODIS, the irregularities of both CLARA-A2 and PATMOS-x during 2001-2003, the lower values of ISCCP and a decline in both PATMOS-x and CLARA-A2 at the end of their time series. The lower COT for PATMOS-x during the period when channel 3a was active on NOAA-16 (i.e. 2001-2003) is not expected and we cannot currently explain it. The liquid REFF case, however, shows a totally different behavior, with systematically lower values by CLARA-A2, higher by PATMOS-x and intermediate values by MODIS and ISCCP. In the 2001-2003 period CLARA-A2 REFF increases considerably as a result of the different shortwave infrared channel being used.

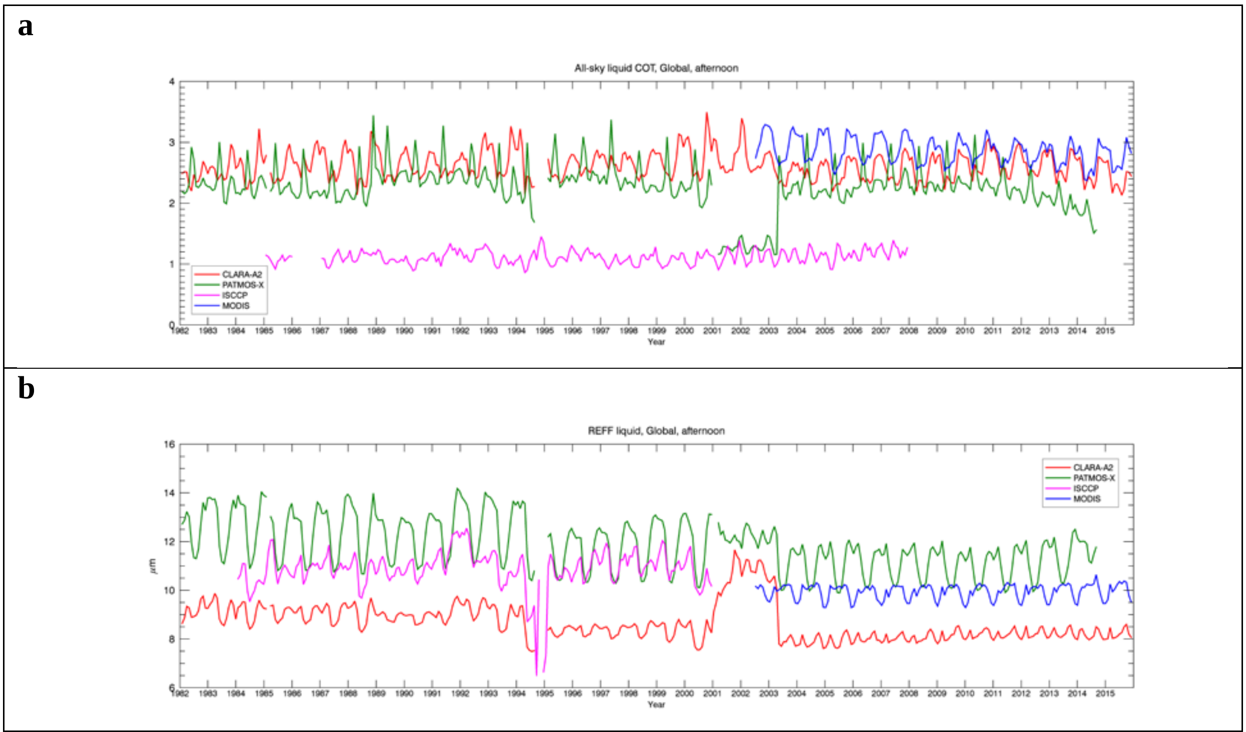


Figure 6.50 *Time series of the afternoon globally averaged all-sky liquid COT from CLARA-A2, PATMOS-x, MODIS and ISCCP (a), and corresponding results for liquid REFF (b).*

6.2.5.7 Evaluation against University of Wisconsin (UWisc) MW LWP data record

The UWisc LWP data record (see Section 5.7) comprises monthly mean all-sky LWP in 1° × 1° grid boxes that is based on all available data for a specific month. In addition, for each month and each grid box over the 1988-2012 period the mean diurnal cycle of LWP is available. In order to obtain the monthly mean all-sky LWP from UWisc closest to the overpass times of the respective NOAA satellites, the mean diurnal cycle parameters, available in the data record, were used to adjust the monthly mean grid box values, based on the equation:

$$\langle LWP(Y,t)\rangle=\langle LWP(Y)\rangle+A\_1 \cos\omega(t-T\_1)+A\_2 \cos2\omega(t-T\_2)$$

where ⟨LWP(Y)⟩ represents the uncorrected monthly mean LWP for year Y, t the local time (h), ω the radial frequency that corresponds to a 24-hour period, and A_1 (T_1) and A_2 (T_2) are the amplitudes (phases) of the first and second harmonics of the diurnal cycle, respectively (see also O'Dell et al. 2008). Since a CLARA-A2 monthly average value may be a composite from multiple satellites, especially in the last years of the time series, a similar procedure was used for the computation of UWisc monthly averages. Specifically, if more than one satellites were available, corresponding UWisc LWP were first computed for each overpass time and then averaged, as in the CLARA-A2 case.

Because microwave instruments are able to penetrate through deep convective clouds or ice over water clouds and measure the LWP at lower altitudes, which is not possible for passive imagers, the present evaluation was restricted to regions with very few (<5%) ice clouds. Therefore, three well-known areas dominated by stratocumulus clouds were selected: the oceanic area west of Africa at 5°-25°S, 10°W-15°E, the area west of South America at 8°-28°S, 70°-90°W, and the area west of California at 20°-30°N, 120°-130°W (see Figure 6.51 for their locations). Obviously, only the ocean parts of these areas are considered for the comparisons because the UWisc data record is restricted to oceans.

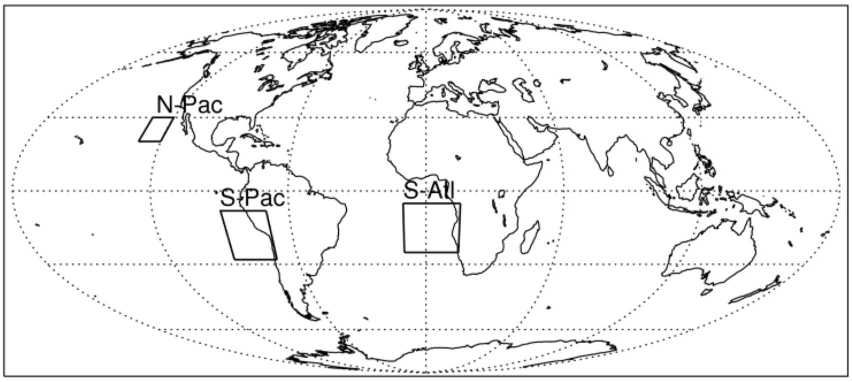
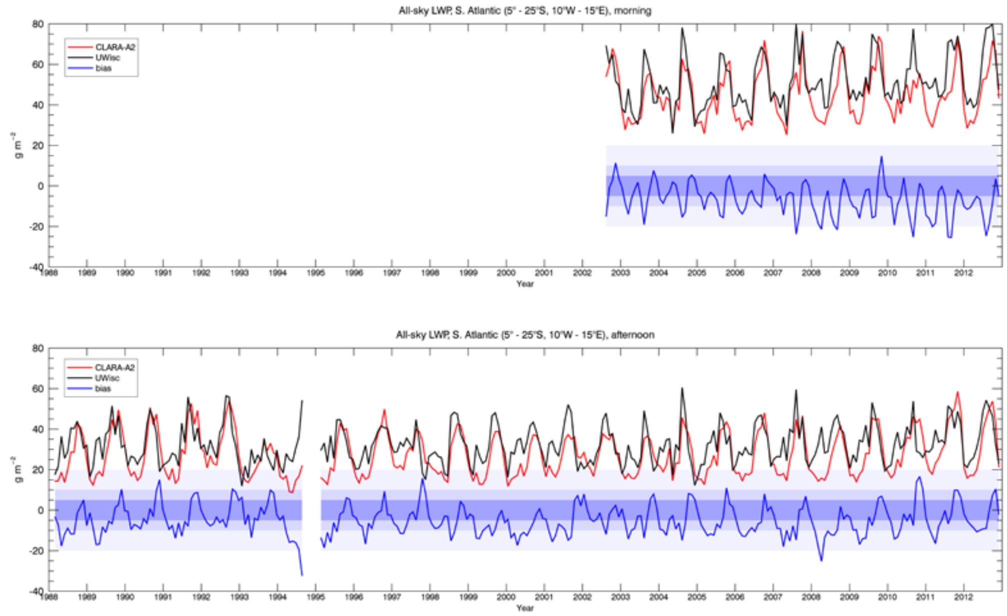


Figure 6.51 *The locations of the South Atlantic (S-Atl), South Pacific (S-Pac) and North Pacific (N-Pac) validation areas.*

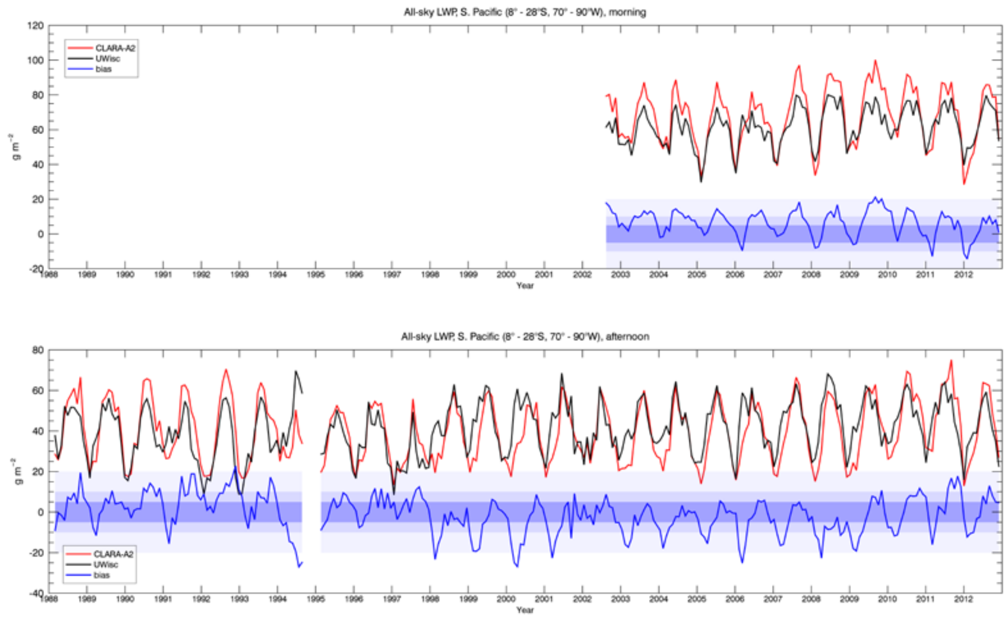
Figure 6.52 shows the time series of the monthly mean all-sky LWP from CLARA-A2 and UWisc, separately for the three areas examined (S. Atlantic, S. Pacific and N. Pacific) and for morning and afternoon satellites, along with the corresponding biases. It is apparent that in all cases the two data records correlate well, with bias fluctuations rarely exceeding the threshold limit. In all areas the morning LWP is consistently higher than in the afternoon in both data records, demonstrating that the general thinning of stratocumulus decks during daytime is well captured. The bias and bc-RMS values reported in Table 6.20 were estimated by combining the data records from the three areas.

a



All-sky LWP, S. Atlantic (5° - 25°S, 10°W - 15°E), morning
All-sky LWP, S. Atlantic (5° - 25°S, 10°W - 15°E), afternoon
CLARA-A2
UWisc
bias
g m⁻²
Year

b



All-sky LWP, S. Pacific (8° - 28°S, 70° - 90°W), morning
All-sky LWP, S. Pacific (8° - 28°S, 70° - 90°W), afternoon
CLARA-A2
UWisc
bias
g m⁻²
Year

c

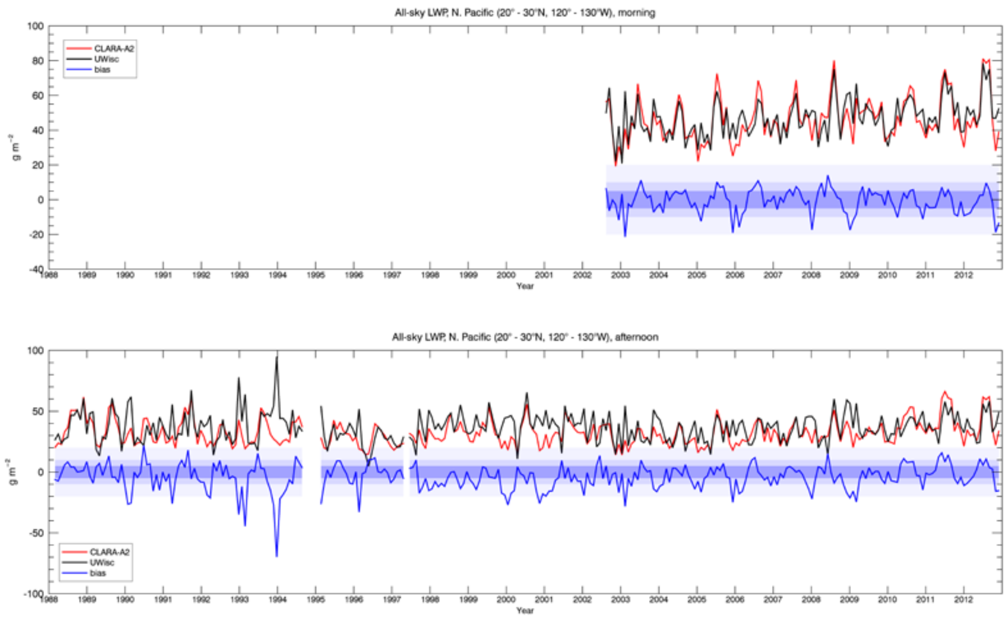


Figure 6.52 *Time series of the monthly mean all-sky LWP from CLARA-A2 and UWisc for the period 1988-2012, over the southern Atlantic (a), the southern Pacific (b) and the northern Pacific (c), separately for morning and afternoon satellites. Corresponding biases are also shown. The shaded areas denote the optimal, target and threshold accuracies for the bias (dark, middle and light, respectively).*

6.2.5.8 Summary of overall LWP validation results

Based on the previously described individual studies of the performance of CLARA-A2 LWP product, Table 6.20 summarizes the results for bias (Mean Error) and bc-RMS. Corresponding compliances with requirements are indicated by YES and NO statements.

Table 6.20 *Overall requirement compliance of the CLARA-A2 all-sky LWP product with respect to the Mean Error and the bias-corrected RMS (bc-RMS). Consistency checks marked in blue. Units are in g m^{-2}.*

Reference data set	Mean Error (Morning/Afternoon)	Fulfilling Threshold requirements (20)	Fulfilling Target Requirements (10)	Fulfilling Optimal Requirements (5)
UWisc	-0.47/-3.43	YES/YES	YES/YES	YES/YES
PATMOS-x	--/4.26	--/YES	--/YES	--/YES
MODIS	3.27/-2.30	YES/YES	YES/YES	YES/YES
ISCCP	17.07/10.37	YES/YES	NO/NO	NO/NO
	bc-RMS	Fulfilling Threshold requirements (40)	Fulfilling Target Requirements (20)	Fulfilling Optimal Requirements (10)
UWisc	10.91/20.09	YES/YES	YES/NO	NO/NO
PATMOS-x	--/17.24	--/YES	--/YES	--/NO
MODIS	11.36/8.20	YES/YES	YES/YES	NO/YES
ISCCP	19.27/14.14	YES/YES	YES/YES	NO/NO

The product generally fulfills target requirements, with the exceptions of the bias when compared with ISCCP and the bc-RMS with respect to UWisc data. It also fulfills optimal requirements for mean error with respect to UWisc, PATMOS-x and MODIS.

6.2.5.9 Evaluation of IWP against MODIS and ISCCP

Figure 6.53 shows the spatial distribution of the 2003-2007 average all-sky IWP, separately from morning and afternoon satellites, from CLARA-A2, MODIS and ISCCP. It should be noted that, unlike the all-sky LWP, PATMOS-x was excluded from this evaluation due to lack of corresponding ice REFF data which would allow a more in-depth inter-comparison. Some major features are similar in all data sets, including the high values of all-sky IWP at the ITCZ, the west Pacific and the Southern Ocean, while larger differences occur near polar regions, where CLARA-A2 acquires higher values in particular over Greenland and Antarctica; these high values should probably be attributed to retrieval issues occurring over ice-covered surfaces. The AVHRR channels are insufficient to distinguish between bright snow/ice surfaces and clouds. The many additional channels on MODIS relative to AVHRR allow a better retrieval of cloud properties over these surfaces.

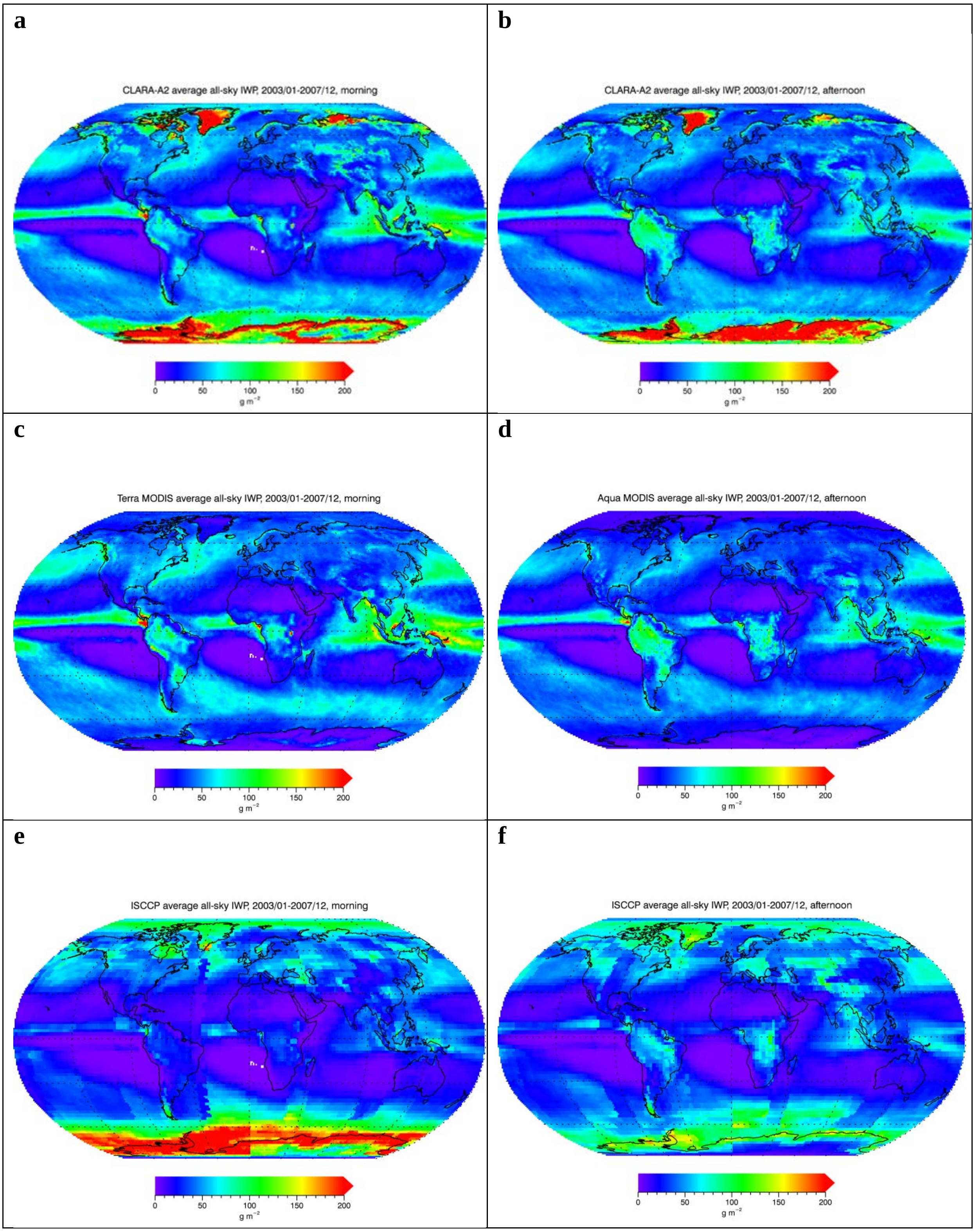


Figure 6.53 *Spatial distribution of the all-sky IWP from CLARA-A2 (a, b), MODIS (c, d) and ISCCP (e, f), separately for morning (left column) and afternoon (right column) satellites, averaged over the period when all data records were available (01/2003-12/2007).*

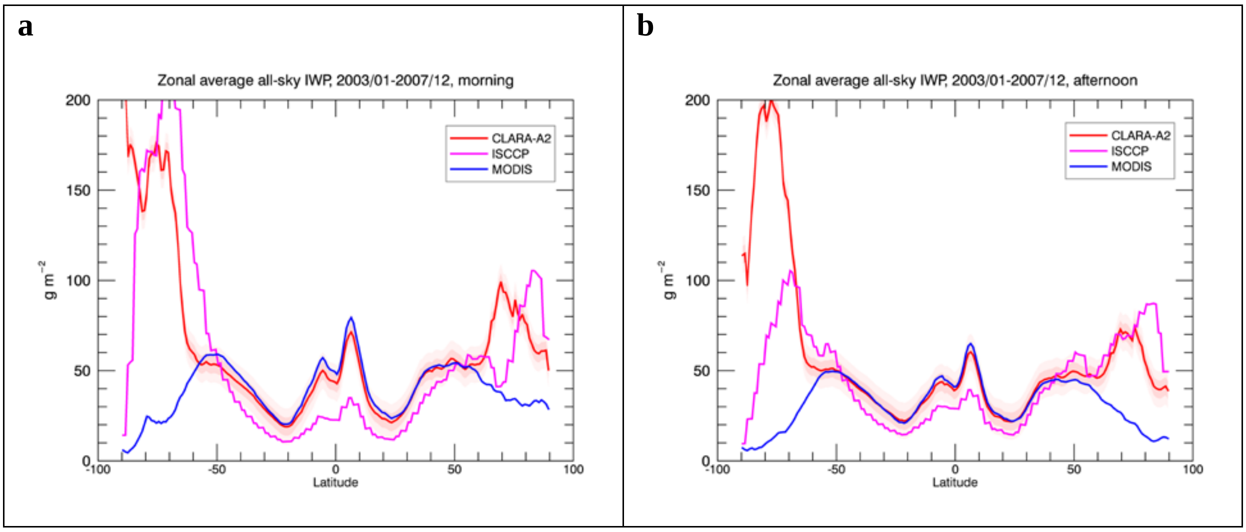


Figure 6.54 *Zonal average all-sky IWP for morning (a) and afternoon (b) satellites, for CLARA-A2, MODIS and ISCCP, computed from corresponding averages from their common period (01/2003-12/2007). The darker and lighter shaded areas around the CLARA-A2 curves denote the optimal and target accuracies, respectively.*

Zonally averaged all-sky IWP (Figure 6.54) from CLARA-A2 and MODIS are in very good agreement between 50°S and 50°N in both morning and afternoon data sets. In the same zone, ISCCP acquires lower values which become higher towards the Polar Regions. Comparing with DARDAR IWP zonal distributions (Eliasson et al., 2011), MODIS IWP appears to be somewhat too low at high latitudes, while CLARA-A2 is too high due to the retrieval issues over bright surfaces as discussed before.

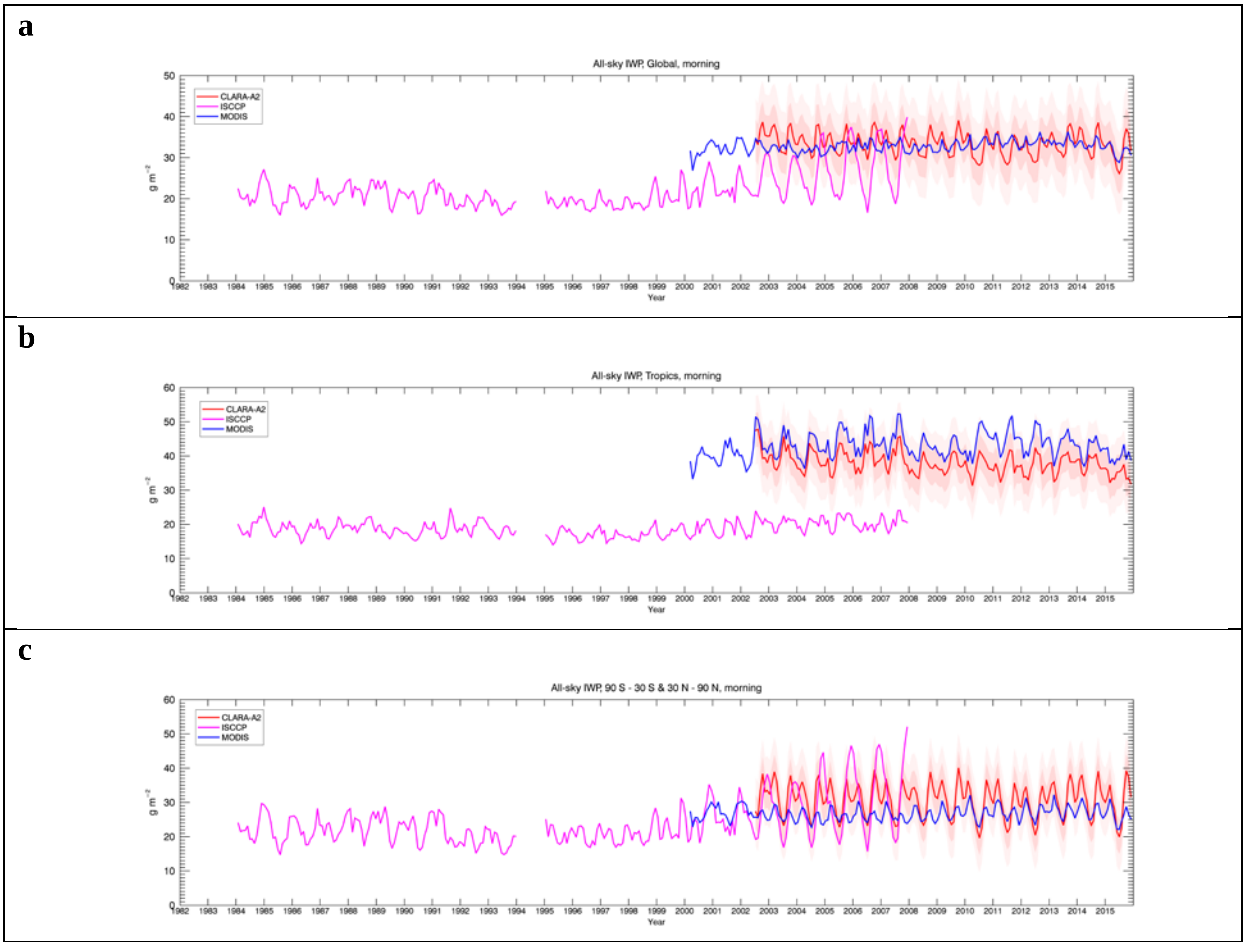


Figure 6.55 *Time series of the morning all-sky IWP from CLARA-A2, MODIS and ISCCP, averaged over the globe (a), the tropics (b) and the areas excluding the tropics (c). The darker and lighter shaded areas around the CLARA-A2 curves denote the optimal and target accuracies, respectively.*

The monthly averaged time series of morning all-sky IWP verifies the good agreement between CLARA-A2 and MODIS globally (Figure 6.55a), although the seasonal cycle is more pronounced in CLARA-A2. This occurs due to the slightly higher MODIS values in the tropics (Figure 6.55b), which are compensated by lower IWP in higher latitudes (Figure 6.55c).

In the afternoon all-sky IWP time series (Figure 6.56), CLARA-A2 is in very good agreement with MODIS in the tropics (Figure 6.56b). At higher latitudes, and consequently on a global scale, CLARA-A2 acquires higher values, while the seasonal cycle of MODIS appears anti-correlated. This pattern should probably be attributed to large differences between the two data sets in Polar Regions. Compared to ISCCP, the agreement is better in the Polar Regions (Figure 6.56c), while in all cases ISCCP values are lower than CLARA-A2. Some irregularities in the CLARA-A2 time series, appearing until 2001, should be attributed to orbital drift issues.

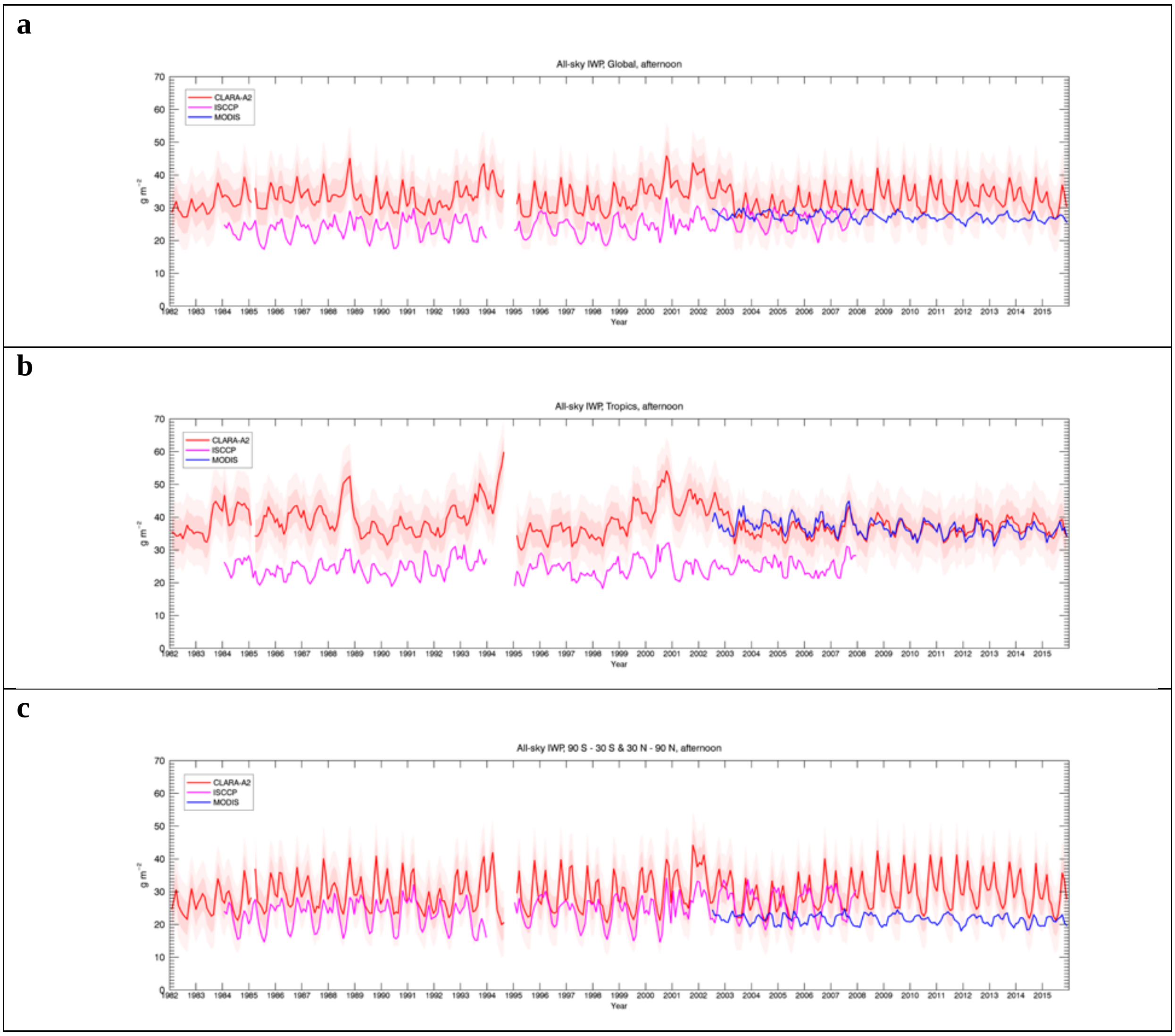
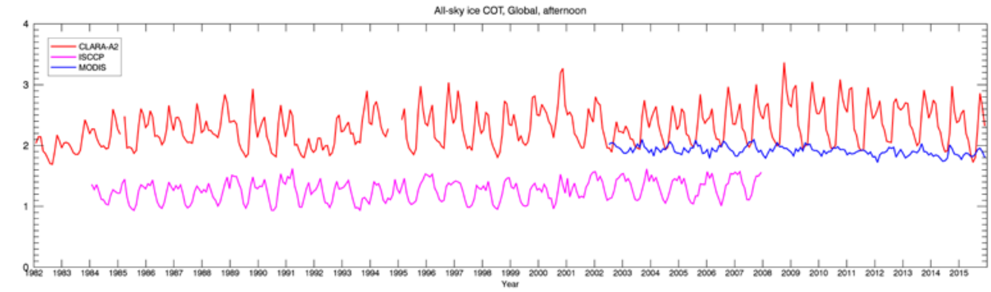


Figure 6.56 *As in Figure 6.55 but for the afternoon satellites.*

6.2.5.10 Evaluation of ice Cloud Optical Thickness (COT) and Effective radius (REFF)

As in the LWP case, time series of all-sky COT and REFF of ice clouds were estimated to further investigate corresponding all-sky IWP results. These time series for the global average case are shown in Figure 6.57. It is apparent that, as in the all-sky IWP case, all-sky ice COT from CLARA-A2 acquires higher values than MODIS, and has a different and more pronounced seasonality (Figure 6.57a). In the ice REFF case, however, there are systematic differences between CLARA-A2 with ISCCP and MODIS, of the order of -5 μm and -2 μm, respectively (Figure 6.57b), while the dip in ISCCP time series in 1995 should probably be attributed to spatial coverage differences with other years.

a



b

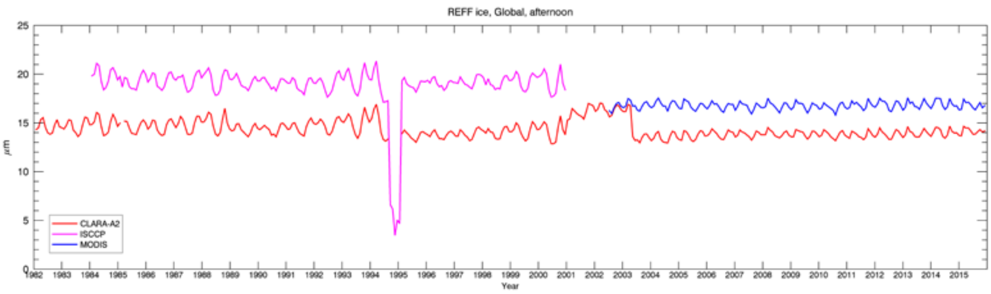


Figure 6.57 *Time series of the afternoon globally averaged all-sky ice COT from CLARA-A2, MODIS and ISCCP (a), and corresponding results for ice REFF (b).*

6.2.5.11 Summary of IWP validation results

Table 6.21 summarizes the results of the CLARA-A2 all-sky IWP comparisons with the other data sets in terms of their mean error and bc-RMS, and their compliance to predefined requirements. Optimal bias requirements are always achieved, while in the bc-RMS case, the target requirement is fulfilled.

Table 6.21 *Overall requirement compliance of the CLARA-A2 all-sky IWP product with respect to the Mean Error and the bias-corrected RMS (bc-RMS). Consistency checks marked in blue. Units are in g m^{-2}.*

Reference data set	**Mean Error (Morning/Afternoon)**	**Fulfilling Threshold requirements (40)**	**Fulfilling Target Requirements (20)**	**Fulfilling Optimal Requirements (10)**
MODIS	0.59/5.09	YES /YES	YES /YES	YES /YES
ISCCP	7.42/8.55	YES /YES	YES /YES	YES /YES
	bc-RMS	**Fulfilling Threshold requirements (80)**	**Fulfilling Target Requirements (40)**	**Fulfilling Optimal Requirements (20)**
MODIS	23.80/19.57	YES/YES	YES/YES	NO/YES
ISCCP	30.69/25.37	YES/YES	YES/YES	NO/NO

6.2.6 Joint Cloud property histograms (JCH)

6.2.6.1 Evaluation against MODIS Collection 6 and PATMOS-x

In this section, CLARA-A2 level-3 JCHs are evaluated against MODIS-Aqua and PATMOS-x JCHs for the overlapping observation years 2003-2014, combining all months. Histograms are compiled on the traditional ISCCP 7x7 CTP-COT bin resolutions, with relative histogram frequencies reported for liquid and ice clouds combined. We choose to display the results with frequencies relative to the 7x7 bins (frequency sums to 100%) rather than to the absolute cloud fraction since we are striving to compare and contrast systematic cloud regime frequencies between CLARA-A2, MODIS, and PATMOS-x, which are identifiable in JCH representations. Only sunlit observations are included in JCH data sets as solar channels are required for the COT retrievals. Since MODIS-Aqua follows and afternoon ascending orbit, only JCHs from CLARA-A2 level-3 satellites overpasses with a local afternoon orbit are analyzed for comparison. Likewise, PATMOS-x JCHs have been computed following the similar JCH processing logic, using PATMOS-x level-2b afternoon orbits and computing monthly distributions. All JCH figures in this section are shown for three data subsets: 1) full grid resolution (sea & land), 2) sea-only grids, and 3) land-only grids; CLARA-A2 JCHs are shown across the top row, MODIS across the middle row, and PATMOS-x across the bottom row.

Global JCHs for all months within the years 2003-2014 are shown in Figure 6.58. Qualitatively, the full global JCH distributions of CLARA-A2 (Figure 6.58a), MODIS (Figure 6.58d), and PATMOS-x (Figure 6.58g), identify global cloud regimes dominated by low-level and upper-level clouds. MODIS. However, these peaks are less pronounced CLARA-A2 compared to the other data sets. Instead CLARA-A2 has a much broader swath of relatively low frequency cloud distributions at mid-levels, spanning the full COT range (Figure 6.58a). This feature is absent in MODIS and PATMOS-x JCHs, where a relative minimum in mid-level JCH frequency is found between the upper- and lower-level cloud peaks (Figure 6.58d, g). This underestimation of cloud top pressure is consistent with the level 2b analysis shown earlier in Figure 6.14 in section 6.1.2.2. The tendency for overestimating mid-level cloud occurrence in CLARA-A2 is also clearly indicated in Figure 6.9 in section 6.1.1.5 (i.e., high cloud CTHs being underestimated and low cloud CTHs being overestimated). It is also evident that a large majority of mid-level cloud overestimation is found over the global seas (Figure 6.58b, e, h). The dominant cloud regimes here are status, stratocumulus and shallow cumulus convection, where MODIS and PATMOS-x cloud tops generally do not penetrate to pressures lower than ~800 hPa. CLARA-A2 tends to overestimate these cloud top heights, and for that reason the maximum cloud frequencies are binned into CTPs ranging 680-800 hPa. The relative minima in MODIS JCH frequency, and especially for PATMOS-x frequency, at mid-levels is more sharply defined for sea-only grids (Figure 6.58e, h), whereas the CLARA-A2 distributions for both sea+land (Figure 6.58a) is clearly related to the enhanced frequency of mid-level cloud top classification over seas (Figure 6.58b). Frequency distributions of the highest clouds tend to be more comparable for CLARA-A2 (Figure 6.58a) and PATMOS-x (Figure 6.58g), than for MODIS (Figure 6.58d).

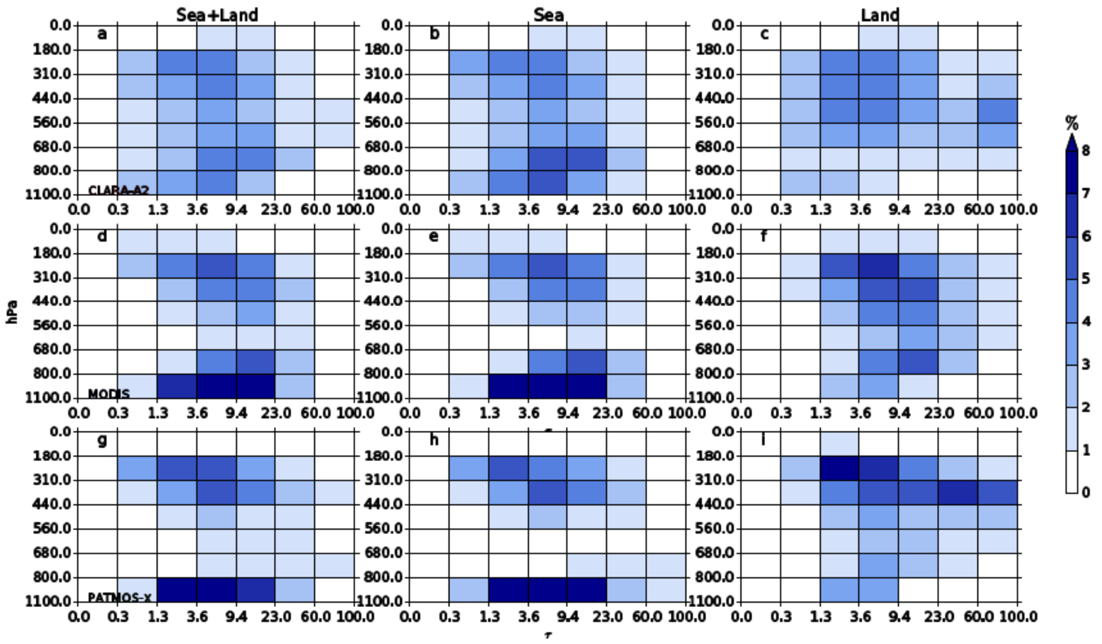


Figure 6.58 *Global JCH relative frequency distributions [colors, %] of CTP [hPa] and COT for all months during 2003-2014. The top row (panels a-c) are CLARA-A2, the middle row (panels d-f) are MODIS Collection 6, and the bottom row (panels g-i) are for PATMOS-x. Left column contains the JCHs over sea and land surfaces (sea+land), middle column over sea-only surfaces (sea) and right column over land-only surfaces (land). Histogram frequencies are normalized to unity, such that each histogram sums to 100%.*

Over land, MODIS and PATMOS-x distributions show an increased frequency of mid- and high-level clouds, and a reduction in shallow cumulus and stratiform clouds (Figure 6.58f, i). A relative increase in very optically thick mid- and upper-level clouds, representative of nimbostratus and deep convection, also emerges for MODIS and PATMOS-x. CLARA-A2 distributions generally agree with this distribution change, although CLARA-A2 tends to observe a higher frequency of optically thinner clouds (COT ranging 0.3-3.6) across the tropospheric column (Figure 6.58c). Furthermore, there is a substantial amount of very optically thick mid- to upper-level clouds in CLARA-A2 and PATMOS-x (Figure 6.58c, i), which are missing entirely in MODIS (Figure 6.58f). In CLARA-A2, this feature is linked to problems in estimating COT properly over snow covered surfaces. A JCH where the Antarctic continent was masked resulted in the removal of this relative peak of high COT as mid- to high cloud levels in CLARA-A2 (not shown).

Regarding the tropical region (30°S-30°N), CLARA-A2 and MODIS JCHs indicate common cloud regime distributions over the tropics (Figure 6.59a, d). There is a tendency for a reduction in both low-level and upper-level COTs, also apparent in PATMOS-x (Figure 6.59g). However, here PATMOS-x differs more relative to the other data sets; essentially only very low and very high clouds are observed (Figure 6.59g). We find that CLARA-A2 generally underestimates the frequency of low-level clouds relative to the other data sets, at the expense of classifying these clouds within the mid-level cloud top range. Again, this is primarily a feature that occurs over the tropical seas, where MODIS and PATMOS-x distinctly indicate a minimum distribution of mid-level clouds (Figure 6.59e, h), while CLARA-A2 retains a substantial mid-level cloud fraction (Figure 6.59b).

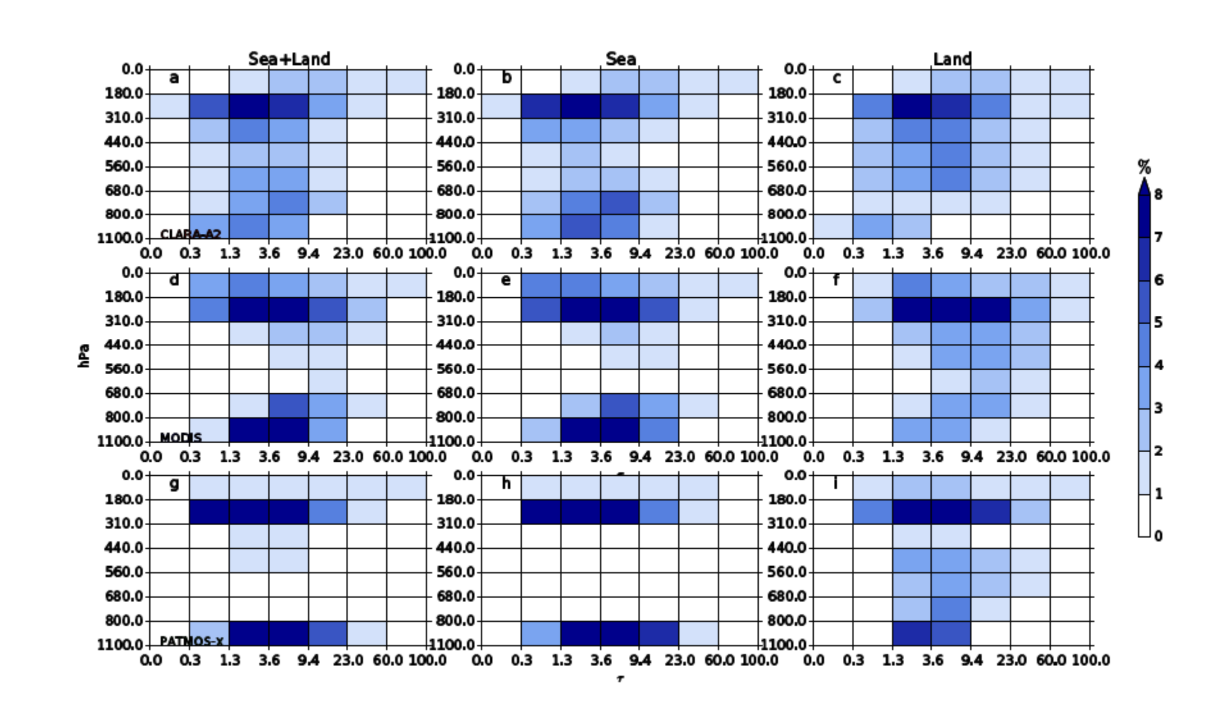


Figure 6.59 *Same as in Figure 6.58, but for the tropics defined as 30°S to 30°N.*

The peak distributions in upper-level, relatively optically thin cirrus show a close agreement between the two data sets for essentially all surface types. Over the tropical region, these high-level cirrus are extremely critical to the top of atmosphere energy budget, and it is crucial that this cloud regime is observed as accurately as possible. Thus the rather strong agreement to MODIS and PATMOS-x is striking. We note that CLARA-A2 distributions generally miss very optically thin clouds that MODIS observes at very high heights, especially over sea (Figure 6.59b, e); PATMOS-x also underestimates the very highest, optically thin clouds (Figure 6.59h) relative to MODIS, but the underestimate is not as large as for CLARA-A2. However it seems that CLARA-A2 is actually observing these low optical thickness clouds, but retrieving their cloud top height from somewhat deeper within the cloud (higher CTP). This may also be related to the CO2 slicing method which MODIS employs to estimate mid- and upper-level cloud top heights (e.g., Pincus et al., 2012).

Over land, JCH distributions show a rather good agreement (Figure 6.59c, f, i). The most apparent difference is an enhanced distribution of low, optically thin clouds for CLARA-A2 that are absent in MODIS and PATMOS-x. If we trust MODIS and PATMOS-x that these clouds are absent, the identification of cloudy scenes is likely related to difficulties in screening pixels for cloud-free cases over highly reflective desert surfaces.

The southern mid-latitude region (30°S-60°S) is dominated by ocean, and as such the JCHs for sea+land (Figure 6.60a, d, g) follow very closely to those from sea-only (Figure 6.60b, e, h). The enhanced occurrence of mid-level clouds, at nearly all COT ranges, is again apparent for CLARA-A2 compared to MODIS and PATMOS-x. This results in less pronounced relative peak frequencies for the low-level and upper-level clouds (Figure 6.60b, e, h). PATMOS-x shows a tendency to further exaggerate the relative minimum frequencies of mid-level clouds compared to MODIS. Generally, the Southern Ocean is contains a high fraction of low-level stratiform cloud cover, which MODIS and PATMOS-x indicate are often with cloud tops at CTP above 800 hPa (Figure 6.60e, h). CLARA-A2 also has a relative maxima in cloud fraction at low-levels (Figure 6.60b), but the cloud tops are often retrieved with a CTP that causes the relative JCH distribution maxima to jump to the next lowest pressure bin (or subsequent CTP bins with lower CTP). The distributions of upper level clouds over sea for all three data sets are in excellent qualitative agreement.

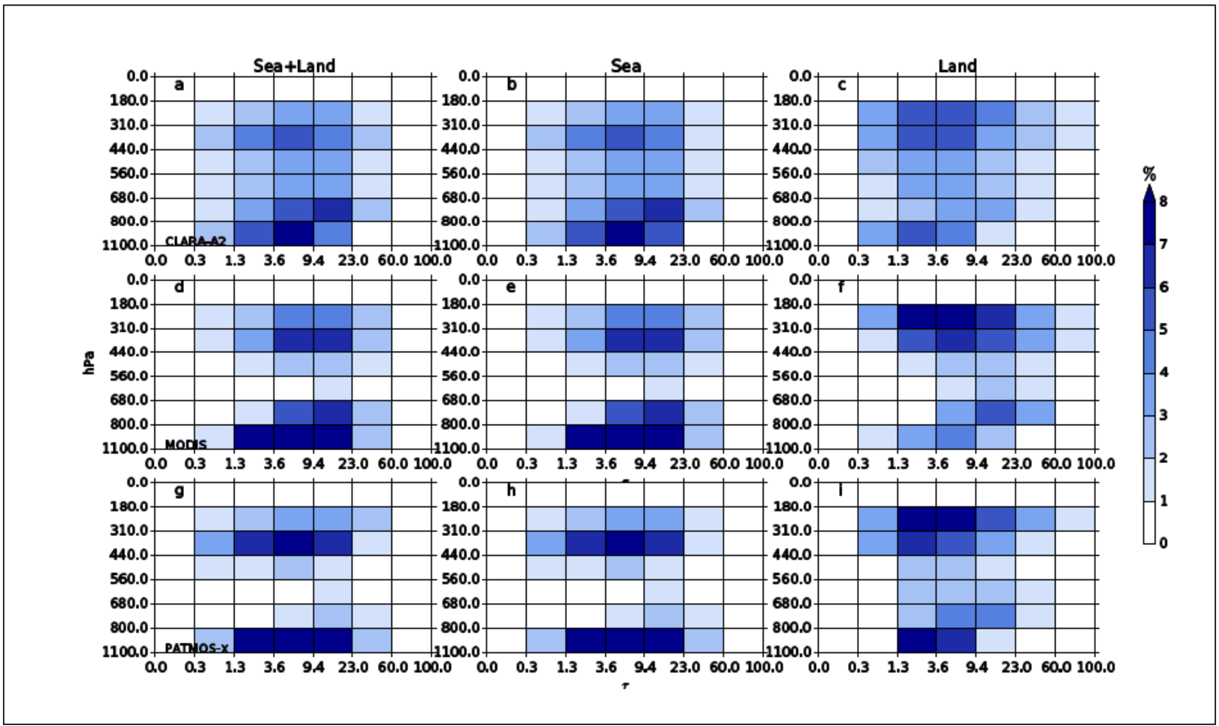


Figure 6.60 *Same as in Figure 6.54, but for the southern hemisphere mid-latitudes defined as 30°S to 60°S.*

Although the contribution of clouds over land in the southern mid-latitudes is relatively small, here JCHs from all data sets show a modest overall distribution agreement. For CLARA and PATMOS-x, the lowest and relatively optically thin clouds show a frequency distribution peak (Figure 6.60c, i), while MODIS identifies a more stratus-type peak at greater COT and lower CTPs (Figure 6.60f). This is likely associated with the biases in CLARA-A2 cloud masking, and subsequent cloud products, over highly reflective surfaces, such as Australia. At upper levels, all data sets indicate a preference for cirrus clouds, but the peak frequencies are higher for MODIS and PATMOS-x. MODIS and PATMOS-x tend to retrieve these clouds at lower CTPs than CLARA-A2, which again may be related to the cloud top retrieval method differences.

For the northern hemisphere mid-latitudes (30°N-60°N), the CLARA-A2 sea+land JCH indicates a rather broad frequency distribution across the troposphere, with modest distribution maxima at low- and upper-levels (Figure 6.61a). These maxima differ slightly from the maxima distributions observed in the MODIS JCH (Figure 6.61d), which show lower, more optically thick low-level clouds, and higher, more optically thick high-level clouds (Figure 6.61d). PATMOS-x distributions show a higher of very low clouds that are optically thinner than both CLARA and MODIS (Figure 6.61g). CLARA-A2 also indicates a subset of clouds with COT ranging 0.3-1.3 across the full troposphere, which is essentially missing from MODIS and PATMOS-x. Over seas, the peak distributions at low and high clouds are more pronounced, but the absolute frequency peaks are still smaller than MODIS and PATMOS-x (Figure 6.61 b, e, h).

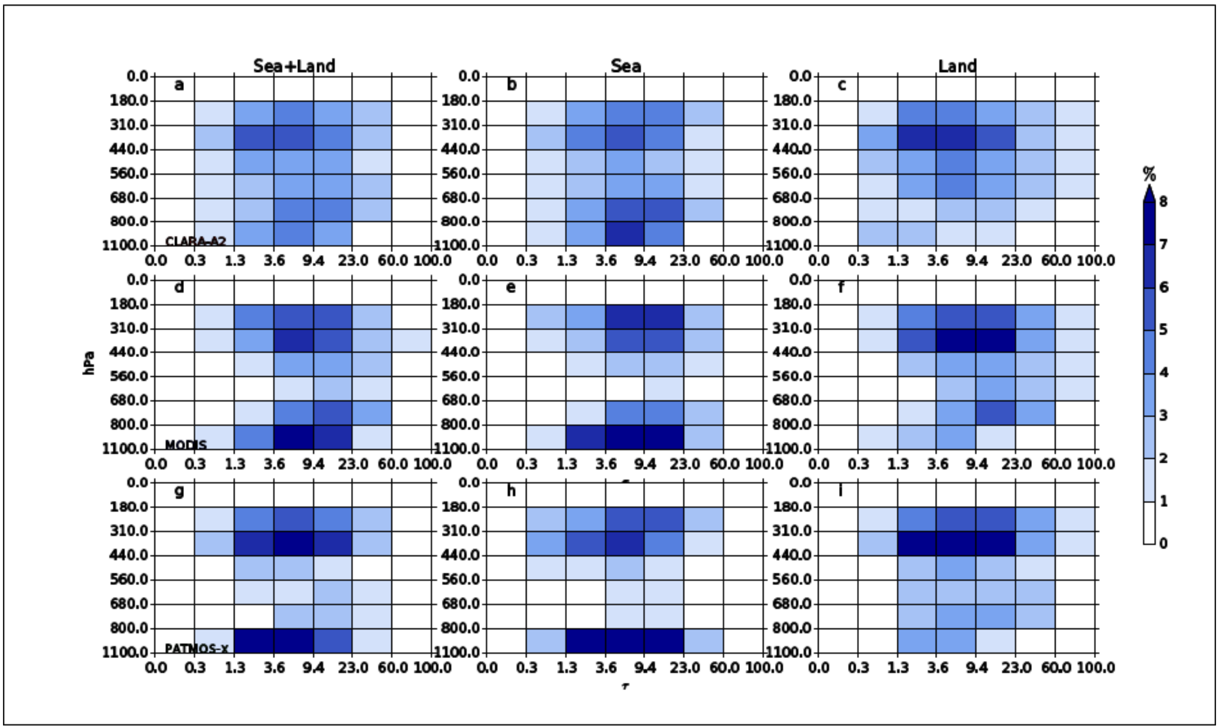


Figure 6.61 *Same as in Figure 6.54, but for the northern hemisphere mid-latitudes defined as 30°N to 60°N.*

Peak distributions in upper-level cirrus or cirro-stratus over land match well between the data sets (Figure 6.61c, f, i). However MODIS and PATMOS-x tend to distribute a larger frequency of these clouds even higher than CLARA-A2. All data sets indicate an increase mid-level cloud frequency over land, but CLARA-A2 tends to have more occurrences of optically thin mid-level clouds that are nearly entirely absent in MODIS and PATMOS-x. MODIS also indicates a relative distribution peak for optically thick low-level clouds (Figure 6.61f) that is generally missing from CLARA-A2 (Figure 6.61c) and only weakly evident in PATMOS-x (Figure 6.61i).

Summary of results

- JCHs provide a unique method of combining multiple data streams to visualize important cloud regime distributions.
- CLARA-A2, MODIS, and PATMOS-x JCHs are qualitatively similar in their cloud regime distributions.
- There is a tendency for relative distribution maxima at low-level (CTP > 680 hPa) and high-level (CTP < 410 hPa). Associated COT peaks range approximately from 1.3 to 23.
- The CLARA-A2 frequency peak mode for low-level clouds is often biased towards lower CTP (higher cloud tops) relative to MODIS and PATMOS-x.

- The CLARA-A2 frequency peak mode for upper-level clouds is often biased towards higher CTP (lower cloud tops) relative to MODIS; the agreement is marginally better for CLARA-A2 and PATMOS-x.
- CLARA-A2 has a tendency for substantially larger fraction of mid-level clouds compared to MODIS and PATMOS-x; these mid-level clouds span nearly the full range of COT.

6.2.7 Process-oriented comparison of products against other data records

In order to rigorously evaluate the CLARA-A2 cloud property data record, it is necessary to investigate the data record through as many different perspectives as possible. The traditional comparisons/validations are not typically tied to any physical process. However, the majority of end users of CLARA-A2 will be studying various processes, climate variability and/or evaluating climate models. Therefore, it is desirable to carry out an initial process-oriented inter-comparison of CLARA-A2 with some reference data sets. This will not only demonstrate how realistically CLARA-A2 represents cloud response to natural variability (and to a particular process), but it will also highlight the nature of the robust response seen commonly in all data sets. To that end, such a comparison was carried out while focusing on three major modes of natural variability; namely, the El Niño Southern Oscillation (ENSO), the Arctic Oscillation (AO) and finally, the Indian Ocean Dipole (IOD). Together these oscillations explain a significant part of the total global natural variability, they affect three major oceanic regions and neighbouring continents and they impact different cloud regimes, thus also providing possibilities to rigorously evaluate CLARA-A2 under different climatic and surface conditions.

Methodology:

- The monthly mean CLARA-A2 cloud property data record was analysed together with monthly mean products from PATMOS-x (version v05r03) and MODIS –Aqua (Collection 6).
- To be consistent with MODIS, the time period for the analysis was restricted from 2003 to 2015.
- The data from only afternoon satellites was evaluated (NOAA-16 from 2003-2005, NOAA-18 from 2005-2008 and NOAA-19 from 2009-2015).
- The periods with enhanced positive and negative phases of the three oscillations in question were selected using time-series of their indices (cf. Figure 6.62).
- The cloud response was calculated in terms of anomalies of cloud properties (cloud fraction and cloud liquid and ice water paths) with respect to climatological means.
- Since the monthly distribution of the positive and negative phases is not uniform (cf. Figure 6.63), the normalized climatological means were calculated to take into account biases in the seasonal distribution of the enhanced oscillation events.
- Due to brevity of space, the evaluation results for exclusively the total cloud fraction are shown in the case of ENSO, AO and IOD, while for LWP and IWP we only show results for the case of ENSO.
- Comparisons for cloud fraction were also done separately for day and night conditions (not shown here).

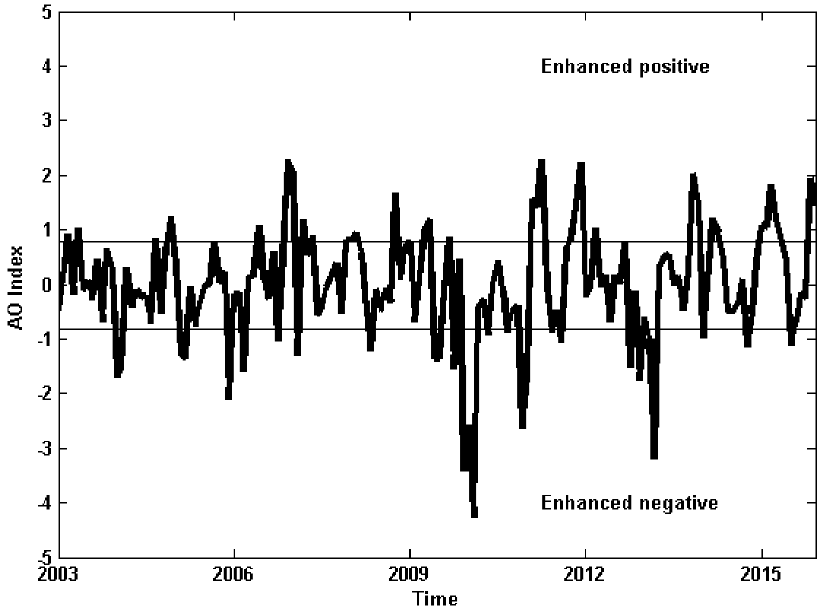


Figure 6.62 *An example of AO index time-series and selected enhanced positive and negative phases of the AO oscillation. All events that exceed (fall below) one standard deviation AO index, shown by thin horizontal line, are considered as enhanced positive (negative) events. Similar criteria were used while selecting events during ENSO and IOD.*

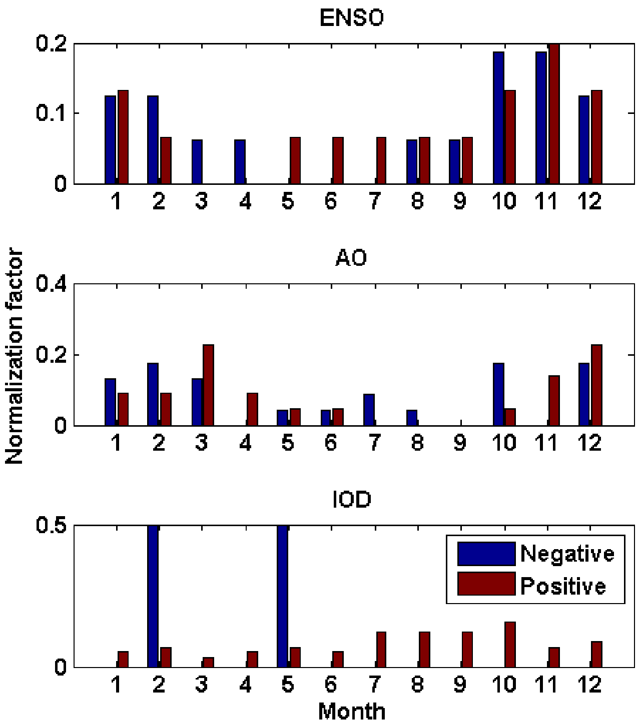


Figure 6.63 *The monthly distribution of enhanced positive and negative oscillation events and the monthly normalization factors used to compute climatological means.*

6.2.7.1 Cloud fraction response to ENSO, AO and IOD

Figures 6.64-6.66 show total cloud fraction anomalies under the enhanced positive and negative phases of the three oscillations in questions and for CLARA-A2, MODIS and PATMOS-x data sets. In Figure 6.64, the pattern correlations of CLARA anomalies with MODIS in the tropics (30N-30S) are 0.98 and 0.97 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.88 and 0.96. The shift in the Walker circulation during positive and negative phases of ENSO is reflected well in all three data sets. The increase (decrease) in cloudiness as a result of weakening (strengthening) of easterly trade winds, decreased (increased) heat transport and increased (decreased) sea-surface temperatures in the eastern Pacific is clearly visible in all data sets. The opposite cloud response in the western Pacific and eastern Indian Ocean is also consistent across all data sets.

Similarly, the high pattern correlations between CLARA-MODIS and CLARA-PATMOS-x are also seen for the IOD events (Figure 6.65). In response to increased (decreased) sea-surface temperatures in the western equatorial Indian Ocean during the positive (negative) phase of the IOD, the total cloudiness shows corresponding increase (decrease). The drought-like conditions in the eastern equatorial Indian Ocean, Indonesia and northern Australia are also captured consistently in all three data sets.
In the case of AO events, the cloud response is somewhat different in the data sets. The pattern correlations of CLARA anomalies with other data sets in the Arctic (60N-90N) are much lower compared to ENSO and IOD events. The correlations with MODIS are 0.58 and 0.72 for the positive and negative phases respectively, and with PATMOS-x the correlations

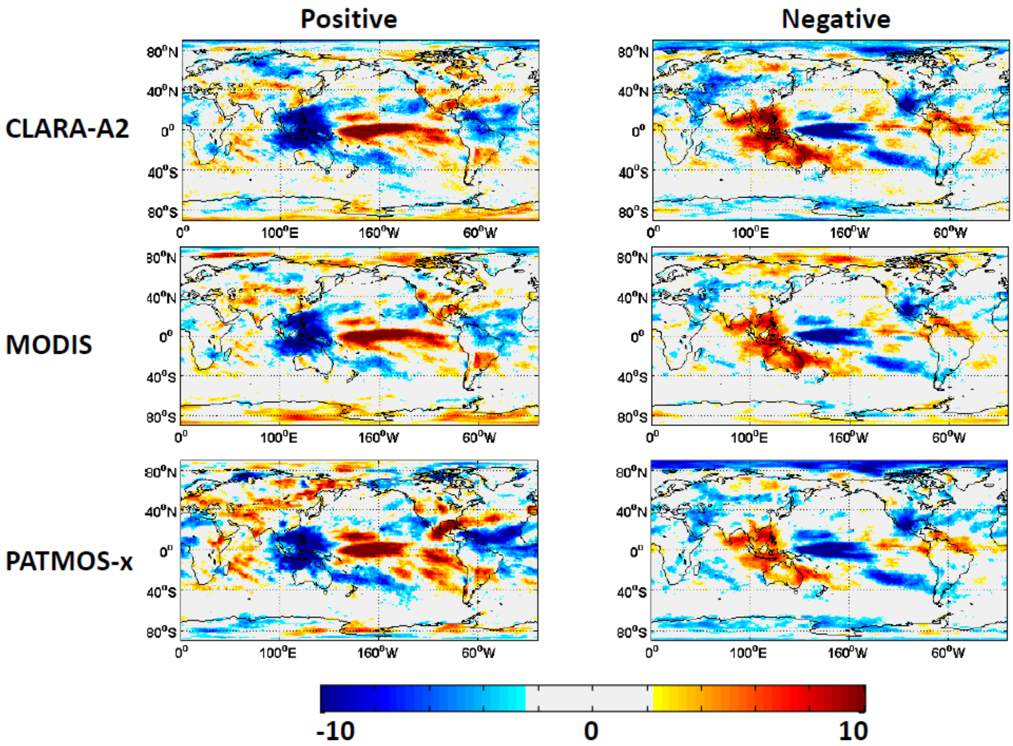


Figure 6.64 *The spatial distribution of total cloud fraction anomalies (in %) observed in three data sets during enhanced positive (strong El Nino) and negative (La Nina) oscillation events. The pattern correlations of CLARA anomalies with MODIS in the tropics (30N-30S) are 0.98 and 0.97 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.88 and 0.96.*

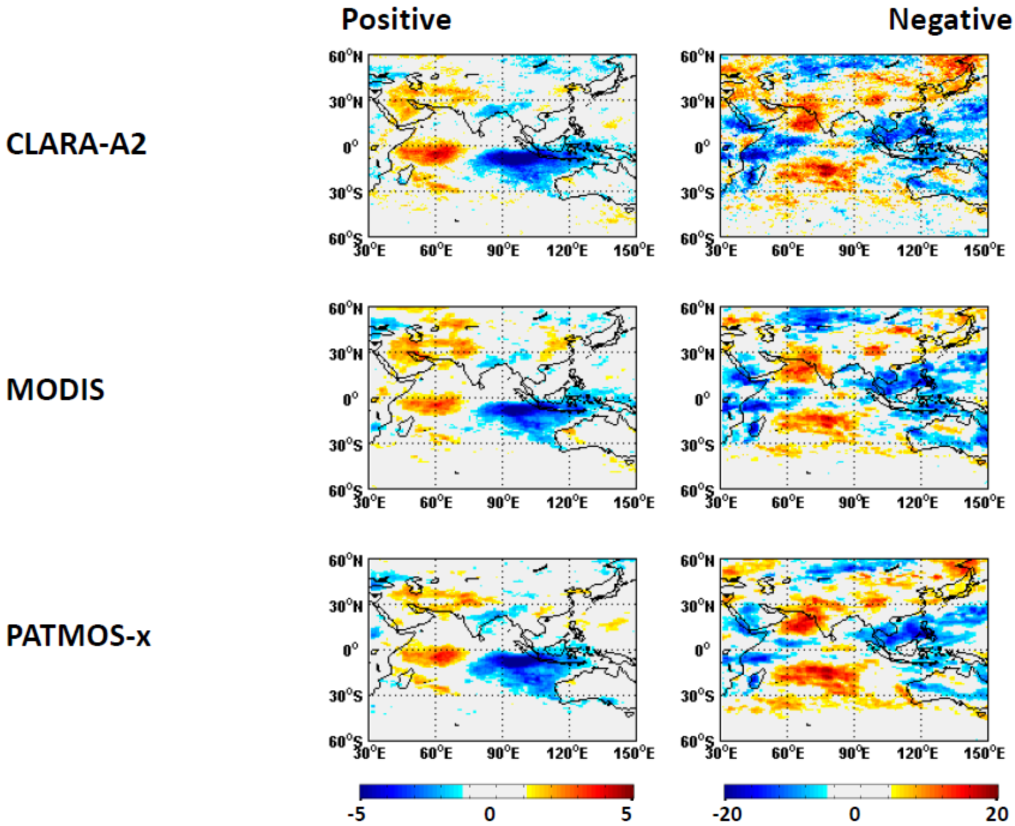


Figure 6.65 *Same as in Figure 6.64, but for the IOD events. The pattern correlations of CLARA anomalies with MODIS in the tropics (30N-30S) are 0.91 and 0.93 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.88 and 0.92.*

are 0.52 and 0.43. This is mainly due to the fact that the enhanced events predominantly occur during polar winter, when the disagreements in the data sets are likely to be strongest. However, it is worth pointing out that the general response in all three data sets is still physical. For example, increase in cloud fraction in the Atlantic sector of the Central Arctic, central Siberia and along the western Norwegian in response to storms reaching the northernmost parts of the Atlantic during the positive phases of AO is seen in all data sets, albeit with different magnitudes. Due to ever persistent cloudiness in the Norwegian Sea and northeastern Atlantic, the changes in cloudiness during the enhanced AO events are not significant.

In general, the cloud response in CLARA-A2 agrees better with MODIS than with PATMOS-x in all cases studied here.

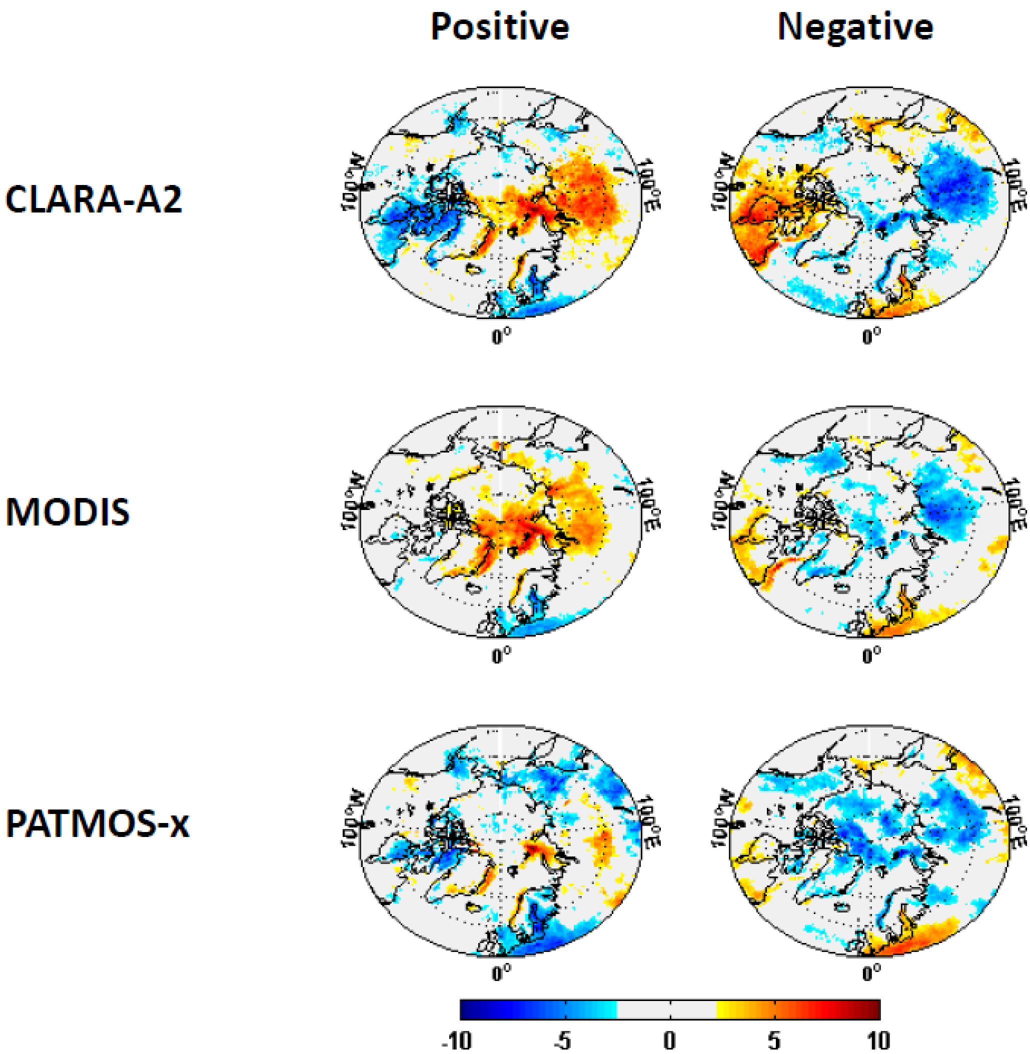


Figure 6.66 *Same as in Figure 6.64, but for the AO events. The pattern correlations of CLARA anomalies with MODIS in the Arctic (60N-90N) are 0.58 and 0.72 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.52 and 0.43.*

6.2.7.2 Changes in cloud condensate during ENSO

Figures 6.67 and 6.68 show anomalies of cloud liquid and ice water paths, respectively, during enhanced ENSO events. Once again, the anomalies of CLARA-A2 LWPs and IWPs are closer to MODIS than to PATMOS-x, as reflected in the values of pattern correlations. For example, both CLARA-A2 and MODIS show bands of increased (decreased) LWP just southward (northward) of the equator in the Pacific during positive ENSO phases, while the reversed anomalies of LWP are observed in both data sets during negative phases. These bands of opposite anomalies occur due to meridional shift in ITCZ and sampling of different cloud regimes. This feature is however either weak or missing in the PATMOS-x data record.

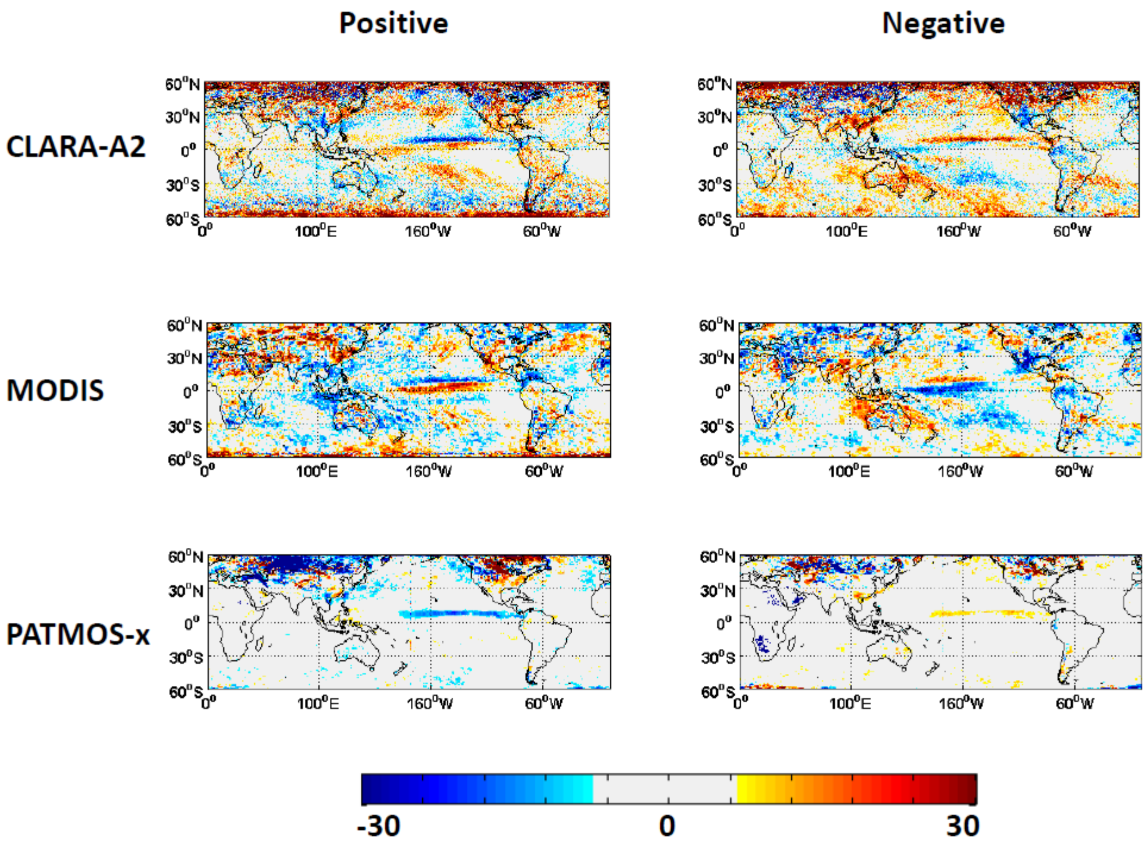


Figure 6.67 *The spatial distribution of LWP anomalies (in g/m2) observed in the three data sets during enhanced positive (strong El Nino) and negative (La Nina) oscillation events. The pattern correlations of CLARA anomalies with MODIS in the tropics (30N-30S) are 0.63 and 0.65 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.36 and 0.22.*

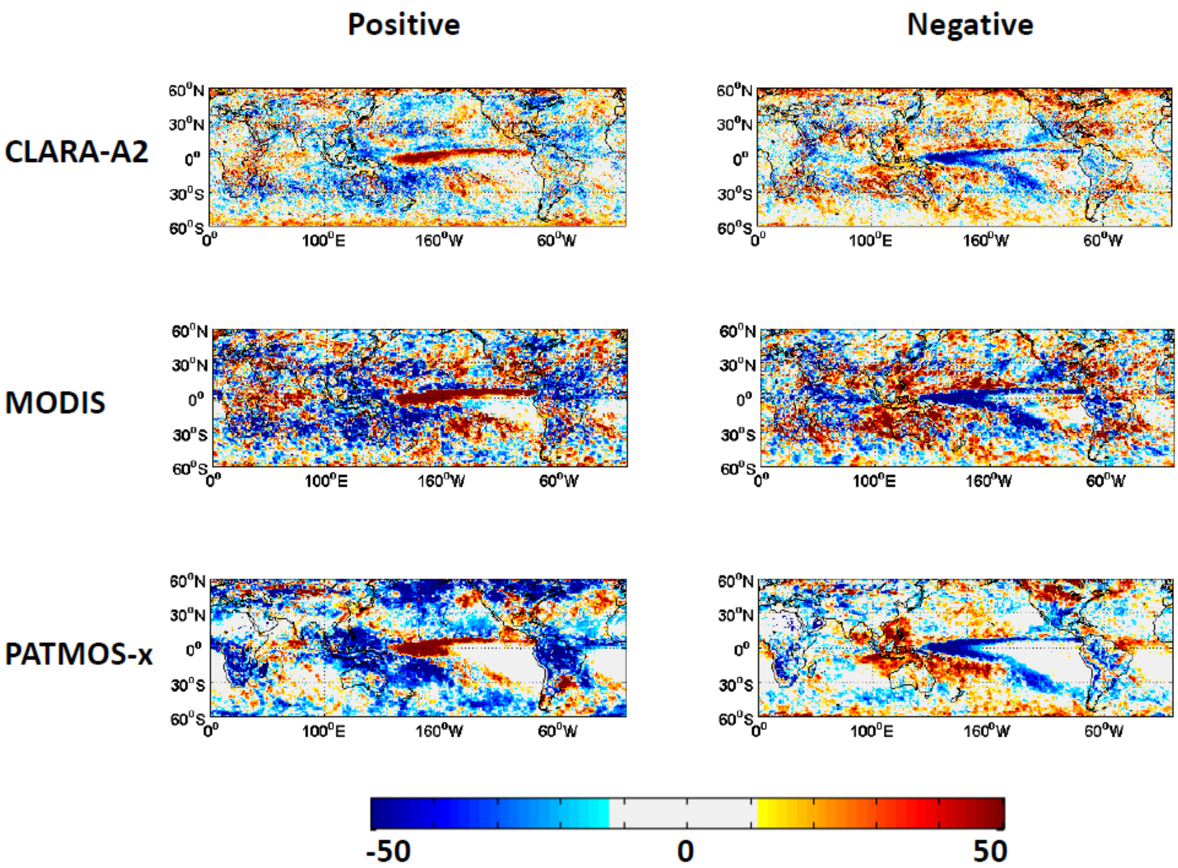


Figure 6.68 *Same as in Figure 6.67, but for IWP anomalies. The pattern correlations of CLARA anomalies with MODIS in the tropics (30N-30S) are 0.79 and 0.82 for the positive and negative phases respectively, and with PATMOS-x the correlations are 0.50 and 0.63.*

6.3 Evaluation of decadal product stabilities

This section covers the evaluation of the decadal stability of CLARA-A2 level-3 products. The evaluation is organized according to Table 6.22.

Table 6.22 *Overview of reference data records used for the evaluation of CLARA-A2 level-3 decadal stability.*

Section	Reference observations	Parameters
6.3.1	SYNOP	CFC
6.3.2	MODIS	CFC, CTP
6.3.3	PATMOS-x	CFC
6.3.4	MODIS	CPH, LWP, IWP

6.3.1 Evaluation of decadal stability against SYNOP observations

The decadal stability gives information on the stability of the data record, for example if the data record has any unnatural trends. To examine the decadal stability, the temporal variation of the bias between the monthly mean cloud fractional cover and the SYNOP monthly mean data record for all available stations is used. Here only the subset of the stations is used, with the constraint that they are available for at least 95 % of the entire time series. This value serves as a good indicator for the stability of the data record. In Figure 6.69 the entire time series of the bias is drawn. The blue line shows the bias and the full variability over the seasonal cycle. The red line is the calculated linear fit. The fit has a decreasing trend of 4.42 % over the entire time series. This gives a decadal trend of 1.30 %. But, as mentioned in section 6.2.1 this is still highly influenced by the number of AVHRRs, which increases with time. This has a strong impact on the representation of the diurnal cycle in the CLARA-A2 data record. This effect is also seen in the bias time series which stabilises after 2001, when the number of simultaneously available satellites gets higher (four or higher).

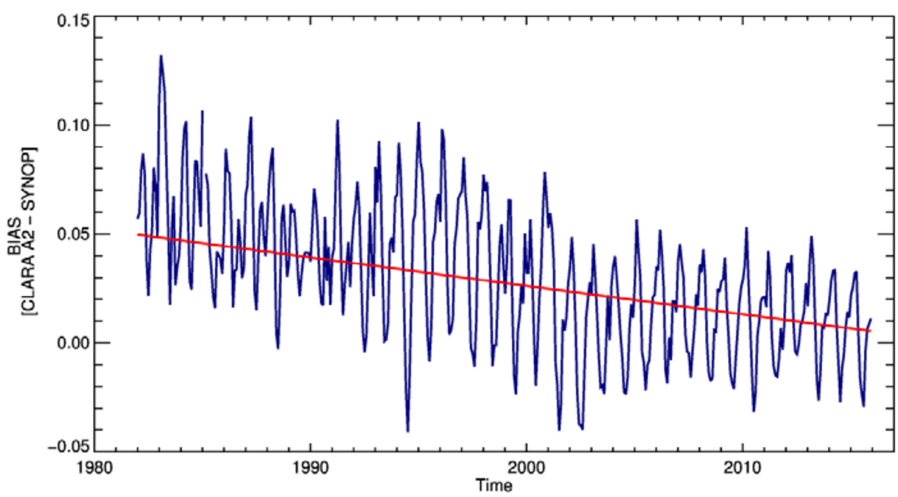


Figure 6.69 *The time series of the bias between the CLARA A2 and the SYNOP cloud fractional cover monthly mean. The red line is the linear fit.*

6.3.2 Evaluation of decadal stability against MODIS observations

6.3.2.1 Fractional Cloud Cover (CFC)

The stability of CLARA-A2 level-3 CFC is evaluated using MODIS (Collection 6) as a nearly independent reference. The 12-yr data (2003-2014) from MODIS-Aqua is used for this purpose. To be consistent with MODIS, the corresponding CLARA-A2 data from NOAA-16 (2003-05), NOAA-18 (2006-09) and NOAA-19 (2010-2014) are used for evaluation. Figure 6.70 below shows the results of the evaluations for the total cloud fraction for different regions globally. It can be seen that, globally, CLARA-A2 CFC clearly achieves the target requirement and the data record is close to satisfying the optimal stability requirement as well. The most robust stability is achieved for the tropical regions, where significant improvements have occurred compared to CLARA-A1. The stability rate is well within the optimal requirement in the tropics. Both southern and northern mid-latitude regions also satisfy optimal requirements. The Polar Regions, however, satisfy only threshold requirements, primarily due to reasons mentioned in 6.2.7.1.

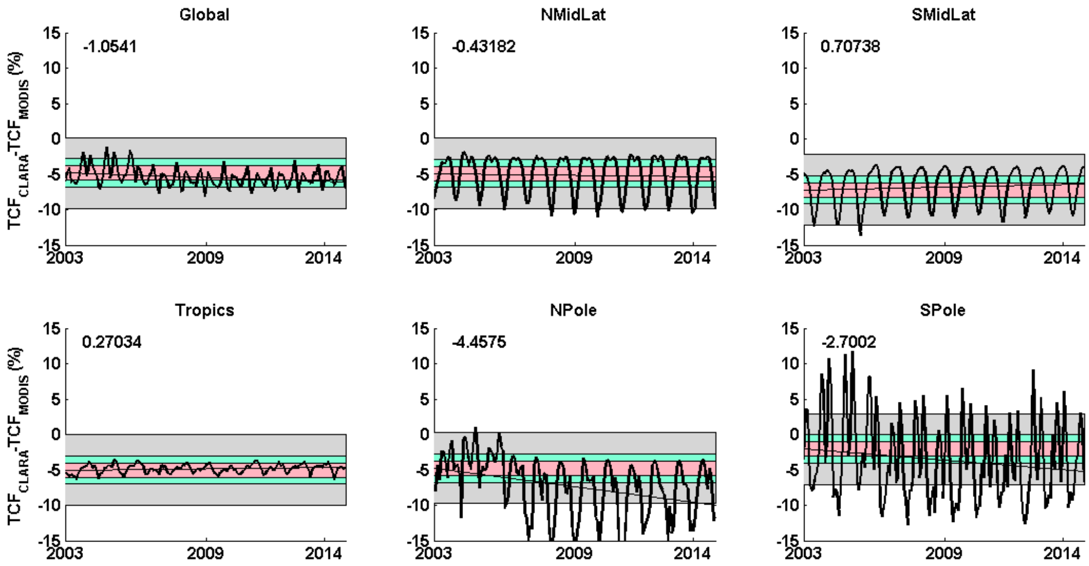


Figure 6.70 *The monthly mean bias in total cloud fraction (CLARA-A2 minus MODIS C6) from 2003 till the end of 2014 for different regions across the globe. The Polar Regions contain areas with latitudes higher than 60 ^{0}, mid-latitude regions are between 30 ^{0}-60 ^{0} and the tropics 30 ^{0}S-30 ^{0}N. The grey, green and pink envelopes show threshold, target and optimal stability requirements respectively. The stability rate (in % per decade) in CLARA-A2 is shown in the top-left corner of each subplot.*

6.3.2.2 Cloud Top level (CTO)

Same as in the case of stability evaluation for CFC, CLARA-A2 and MODIS-Aqua Collection 6 level-3 data are processed for the analysis of cloud top pressure. Figure 6.71 shows the results of the rate of bias stability per decade for cloud top pressure. Globally, CLARA-A2 CTP satisfies the optimal stability requirement. The best results are observed in the case of northern hemispheric mid-latitude regions, while the Polar Regions show strong variability in the bias and its trend. Except the Polar Regions, all other regions across the globe satisfy the optimal stability requirement. The Antarctic region satisfies the target requirement, while the Arctic region satisfies the threshold requirement. However, the inter-annual variability in bias is strong over both Polar Regions.

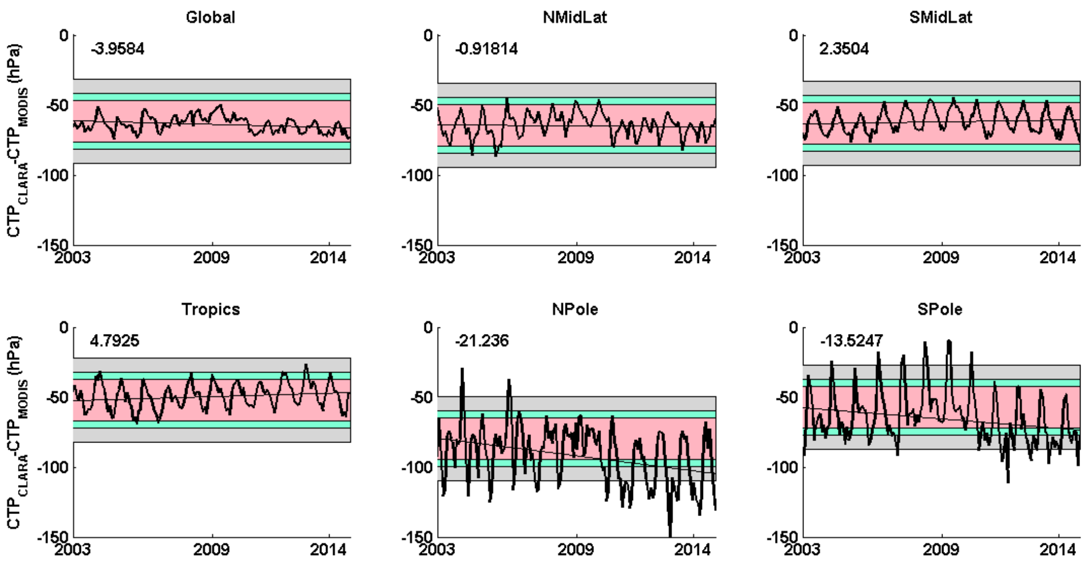


Figure 6.71 *Same as in Figure 6.70, but for cloud top pressure. The grey, green and pink envelopes show threshold, target and optimal stability requirements respectively. The stability rate (in hPa per decade) in CLARA-A2 is shown in the top-left corner of each subplot.*

6.3.3 Evaluation of decadal stability against PATMOS-x

6.3.3.1 Fractional Cloud Cover (CFC)

The stability of CLARA-A2 level-3 CFC is evaluated using PATMOS-x as the reference (or, rather, as a consistency check). Figure 6.72 below shows the results of the evaluations for the total cloud fraction for different regions. For global results the overall trend is small and well within requirements. The agreement is very good over all regions except over the Polar Regions where substantial seasonal differences can be seen (i.e., PATMOX-X cloud amounts are much higher during the Polar Winter).

Nevertheless, it is important to point out that these results do not mean that the two data records agree completely. We have noted considerable differences and even opposite trends for selected periods (see Figure 6.12). Thus, despite using exactly the same AVHRR FCDR there are significant differences which may come from either different methodologies or differences in the selection (i.e., quality control procedures) and sampling of data.

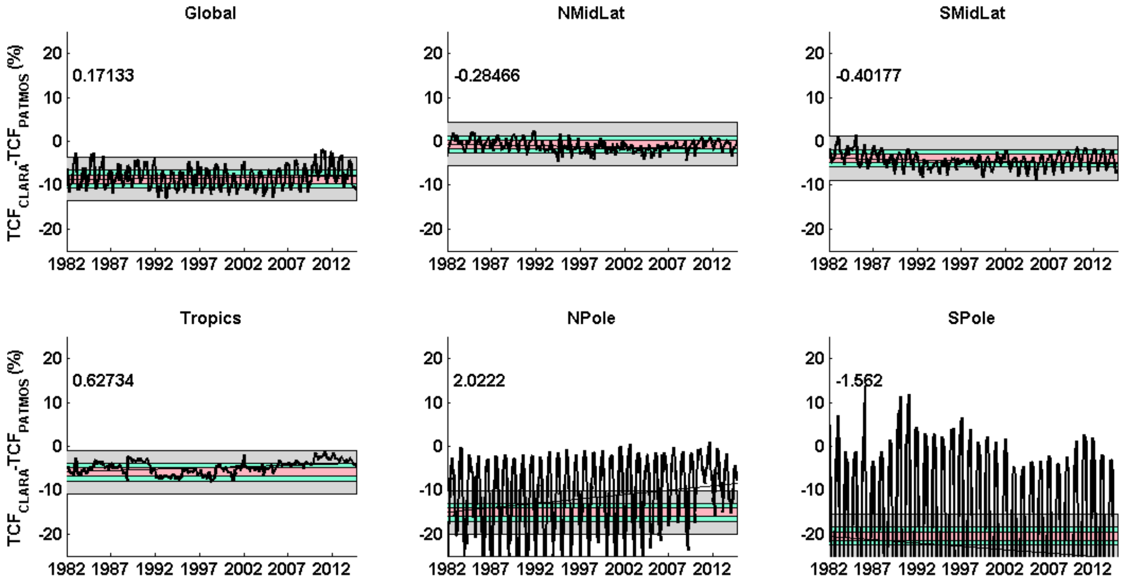


Figure 6.72 *The monthly mean bias in total cloud fraction (CLARA-A2 minus PATMOS-x) from 1982 till the end of 2014 for different regions across the globe. The Polar Regions contain areas with latitudes higher than 60 ^{o}, mid-latitude regions are between 30 ^{o}-60 ^{o} and the tropics 30 ^{o}S-30 ^{o}N. The grey, green and pink envelopes show threshold, target and optimal stability requirements respectively. The stability rate (in % per decade) in CLARA-A2 is shown in the top-left corner of each subplot.*

6.3.4 Decadal stability CPP products

The decadal stability of CLARA-A2 level-3 CPH, LWP and IWP was evaluated using corresponding MODIS Collection 6 level-3 data products. While other satellite data sets may include inherent variability caused by various factors, which prevents an objective evaluation of CPP products stability, MODIS is considered the most stable sensor in terms of calibration and orbital drift issues. For this evaluation, all possible combinations of annual average bias between CLARA-A2 and MODIS time series larger than 10 years were created and, for each combination, decadal trends were estimated, based on a linear regression approach. These trends were then aggregated into an array, which gives an overview of typical values and ranges of decadal bias trends throughout the entire time series, highlighting cases of deviation from the assumed stability of the time series. Overall decadal stability was assessed based on the estimated trend of the entire bias time series.

6.3.4.1 Cloud Phase (CPH)

Figure 6.73 shows the pattern of decadal trends in CPH (a) and CPH_Day (c) biases for afternoon satellites. These values were computed based on a linear regression approach, from all possible combinations of start and end years spanning at least a decade, as previously. Corresponding time series of the annual average bias values are also shown (Figure 6.73b and d). It should be noted that the morning CPH and CPH_Day stability was not examined here, due to the MODIS Terra degradation issue, which affects these products and prevents a meaningful evaluation. It is apparent, however, that there is a positive trend in the afternoon cases of both CPH and CPH_Day, regardless of the combination of start and end years. As shown earlier in Figures 6.40a and 6.44a, this trend should be attributed to a slight decrease in

Aqua MODIS CPH and CPH_Day values during the last years of the time series, for which we currently do not have an explanation, except for a possible degradation occurring as in the Terra MODIS case.

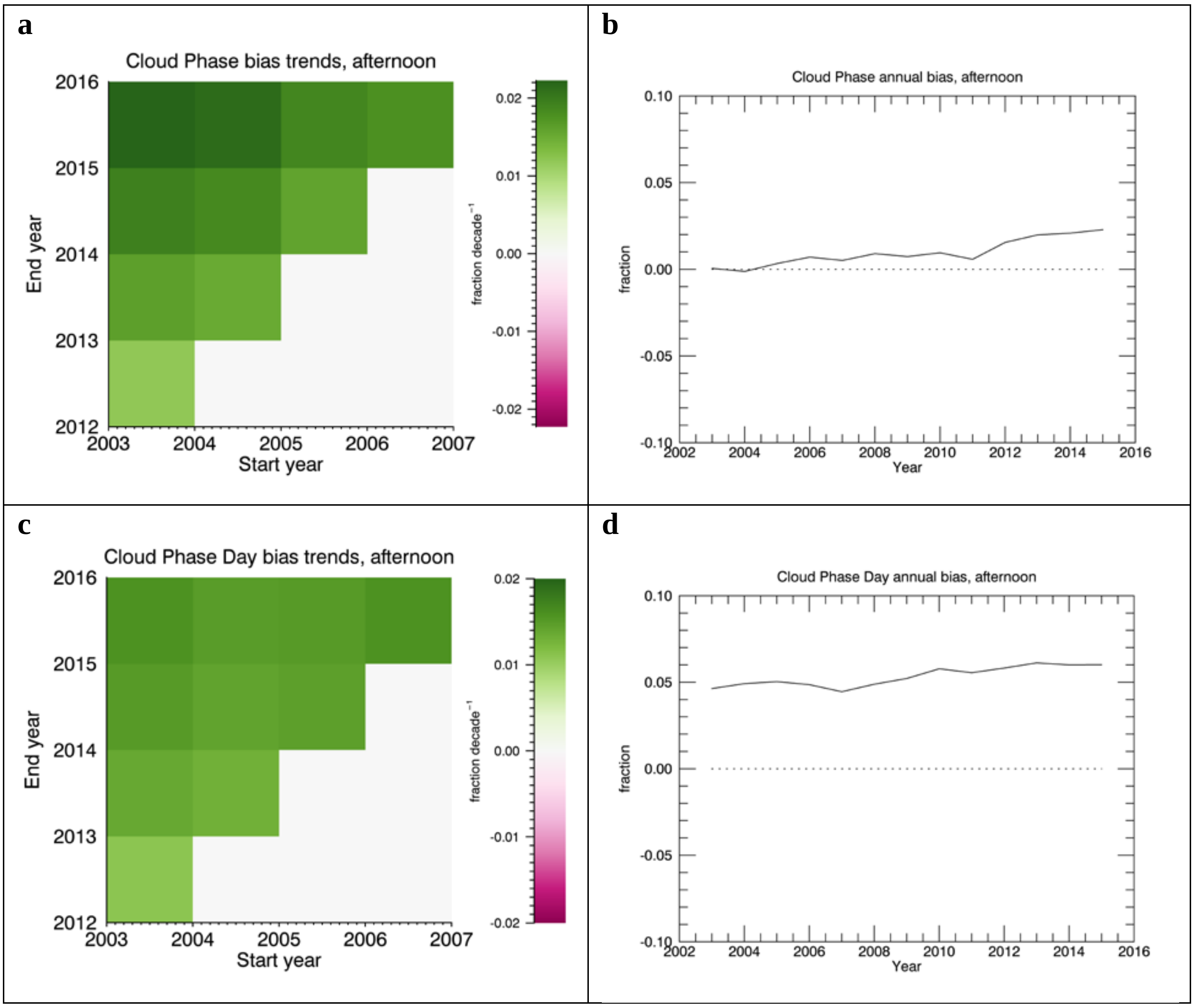


Figure 6.73 *Decadal trends of CPH (a) and CPH_Day (c) bias between CLARA-A2 and MODIS (in fraction decade^{-1}), for afternoon satellites, estimated from all possible combinations of time periods equal or larger than 10 years. Corresponding time series of annual average biases are also shown (b and d).*

Table 6.23 summarizes the results in decadal stability of CPH and CPH_Day biases, and their compliance with the predefined requirements. Bias trends calculated from the entire time period (2003-2015) of data sets were used for this purpose. In the CPH case, decadal stability is slightly over the target requirement, while in the CPH_Day case, target requirement is fulfilled.

Table 6.23 *Overall decadal stability requirement compliance of the Cloud Phase and Cloud Phase Day products. Units are in fraction decade^{-1}.*

CPP product	Trend (Afternoon)	Fulfilling Threshold Requirements (0.05)	Fulfilling Target Requirements (0.02)	Fulfilling Optimal Requirements (0.01)
Cloud Phase	0.022	YES	NO	NO
Cloud Phase Day	0.016	YES	YES	NO

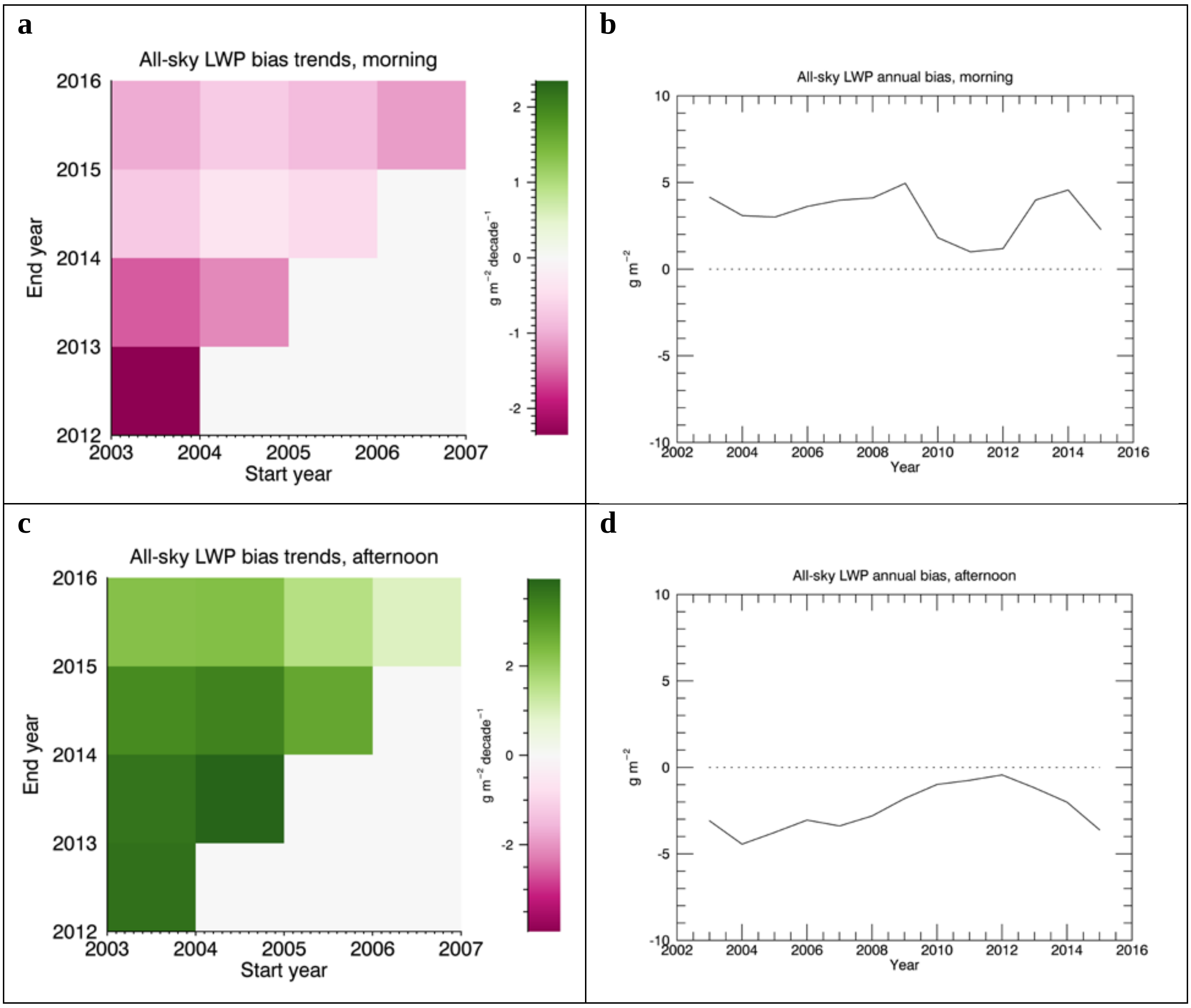


Figure 6.74 *Decadal trends of the all-sky LWP bias between CLARA-A2 and MODIS (in g m^{-2} decade^{-1}), separately from morning (a) and afternoon (c) satellites, estimated from all possible combinations of time periods equal or larger than 10 years. Corresponding time series of annual average biases are also shown (b and d).*

6.3.4.2 *Liquid Water Path (LWP)*

Figure 6.74 shows the bias trend patterns and annual average time series results for morning and afternoon all-sky LWP. The morning time series is quite stable (Figure 6.74b), with negative trends fulfilling the target requirement of 3 g m-2 (Figure 6.74a), due to the relatively low bias values in 2010-2012 and 2015. These deviations coincide with the end of data availability from NOAA-17 (February 2010) and the beginning of MetOp-B data availability in January 2013.

In the afternoon case, positive trends prevail, due to the relatively lower annual bias values at the first years of the time series. When compared against corresponding monthly time series of all-sky LWP (Figure 6.49a), it is apparent that these changes in bias are caused by changes in CLARA-A2, rather than Aqua MODIS.

The compliance of the all-sky LWP stability with the predefined requirement is summarized in Table 6.24. Based on the morning and afternoon trends estimated from the entire time series, it is found that in both cases the target requirement is fulfilled.

Table 6.24 *Overall decadal stability requirement compliance of the all-sky LWP bias against MODIS. Units are in g m^{-2} decade^{-1}.*

Trend (Morning/Afternoon)	Fulfilling Threshold Requirements (5)	Fulfilling Target Requirements (3)	Fulfilling Optimal Requirements (1)
-1.01/2.26	YES/YES	YES/YES	NO/NO

6.3.4.3 *Ice Wather Path (IWP)*

The all-sky IWP trend patterns and time series of annual biases are shown in Figure 6.75. Both morning and afternoon cases are very similar to the corresponding LWP results, with small deviations occurring in 2010-2012 and 2015 in the morning case (Figure 6.75b), leading to negative trends (Figure 6.75a).

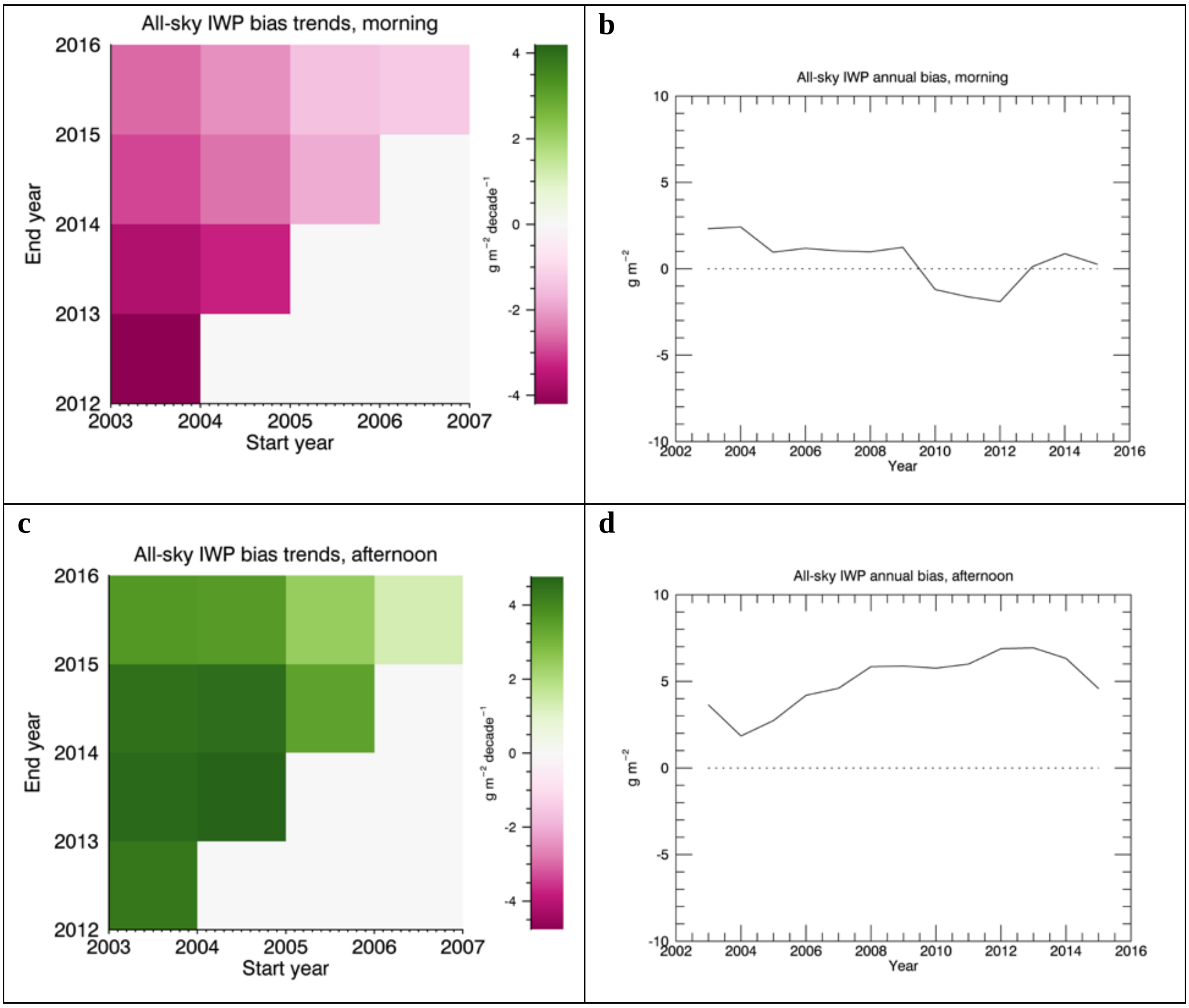


Figure 6.75 *As in Figure 6.74 but for the all-sky IWP.*

In the afternoon all-sky IWP time series (Figure 6.75d), as in the all-sky LWP case, higher biases towards the end of the time series cause positive trends (Figure 6.75c).When examined in combination with the monthly average time series (Figure 6.56a), it becomes obvious that these increased bias values are due to the enhanced seasonality in CLARA-A2, which, however, is not captured by MODIS throughout the time series.

The long term trends in all-sky IWP bias fulfil the target stability requirement in both morning and afternoon cases, as shown in Table 6.25.

Table 6.25 *Overall decadal stability requirement compliance of the CLARA-A2 all-sky IWP product. Units are in g m^{-2} $decade^{-1}$.*

Trend (Morning/Afternoon)	Fulfilling Threshold Requirements (10)	Fulfilling Target Requirements (6)	Fulfilling Optimal Requirements (2)
-2.65/3.67	YES/YES	YES/YES	NO/NO

7 Conclusions

An extensive validation of cloud products from the CM SAF GAC Edition 2 data record has been presented in this report. The reference data records were taken from completely independent and different observation sources (e.g. SYNOP, CALIPSO-CALIOP, SSM/I and AMSR-E) as well as from similar satellite-based data records from passive visible and infrared imagery (MODIS, ISCCP and PATMOS-x). Studies were made based on a mix of level-2 and level-3 products, also addressing some specific aspects affecting inter-comparisons (e.g., cloud detection capabilities for very thin clouds). More in depth inter-comparisons were also made with the PATMOS-x data record because of the close relation (being also based on AVHRR GAC data).

In addition to the larger emphasis on the evaluation of level-2 products compared to the corresponding validation activity for the CM SAF GAC Edition 1 (CLARA-A1), more work has also been spent this time on evaluating Joint Cloud Histograms and the decadal stability. For the latter, the only independent source (having long enough observation capability) is SYNOP reports of cloud cover (CFC). However, we have here added studies based on the MODIS 13-year observational record since we regard MODIS products as high-quality and stable products now having reached a very mature level in the sixth reprocessing effort (i.e., MODIS Collection 6). Similar comparisons have also been made against PATMOS-x for the whole time series but these results have a limited value (due to the high correlation with the CM SAF GAC data record).

Tables 7.1-7.3 below give an overview of all results with respect to the target accuracies, target precisions and requirements on decadal stabilities. How these results were derived and what assumptions and definitions that were used are outlined in detail in the specific sub-sections of this report.

Results show the following, product by product:

- **Fractional Cloud Cover (CFC)**

- The CM SAF GAC CFC product fulfils the Target requirements for both accuracy and precision when compared with all references
- The only exception can be seen for the precision of level-2 products compared with CALIPSO-CALIOP. However, we claim that this is due to an existing mistake in the current requirements (i.e., RMS values should be higher for level-2 products than for Level 3 products).
- The requirement on decadal stability is fulfilled.

- **Cloud Top level (CTO)**

- The CM SAF GAC CTO level-3 product fulfils the Target requirements for all references except against MODIS
- The CM SAF GAC CTO level-2 product fulfils Threshold requirements and is very close to fulfilling also Target requirements.
- The requirement on decadal stability is fulfilled.

- **Cloud Thermodynamic Phase (CPH)**

- The CM SAF GAC CPH product fulfils optimal accuracy requirement against most references except against ISCCP, while the CPH-Day product always fulfils the target requirement.
- In both products, optimal precision requirement is fulfilled against most references except ISCCP, where target requirement is achieved.
- The target and threshold requirements for decadal stability are fulfilled for CPH-Day and CPH, respectively.

- **Liquid Water Path (LWP)**

- The CM SAF GAC LWP product fulfils optimal accuracy and target precision requirements with respect to the UWisc data set. Note that – as a consequence of necessary selections of the data – the validation with UWisc was restricted to oceanic, stratocumulus-dominated areas.
- Optimal accuracy requirement is fulfilled with respect to MODIS and PATMOS-x data records and threshold requirement is achieved with respect to ISCCP. Target precision requirement is achieved with respect to all data sets.
- The target requirement for decadal stability is fulfilled with respect to MODIS.

- **Ice Water Path (IWP)**

- The CM SAF GAC IWP product fulfils optimal accuracy requirements when compared with MODIS and ISCCP.
- Using the same data sets, target precision requirement is achieved.
- The target requirement for decadal stability is fulfilled with respect to MODIS.

- **Joint Cloud property Histograms (JCH)**

- This product is excluded from specific requirement testing because of being composed by two already existing products (COT and CTP)
- Nevertheless, the product has been inter-compared with corresponding results from ISCCP, MODIS and PATMOS-x showing many similarities but also some CLARA-A2 specific features.
- It is believed that the access to this product representation would greatly enhance the usefulness of the CM SAF GAC products in some applications (e.g., in climate model evaluation it is a central product for COSP simulators).

Table 7.1 *Summary of validation results compared to target accuracies for each cloud product. Notice that accuracies are given as Mean errors or Biases (both terms being equivalent) valid for both negative and positive deviations. Results from consistency checks (not totally independent) are marked in blue.*

Product		Accuracy requirement (Mean error or Bias))	Achieved accuracies
Cloud Fractional Cover	(CFC)	5 % (absolute)	-3.2 % (CALIPSO level-2) -3.1 % (SYNOP level-3) -4.9 % (PATMOS-x level-2b) -4 % (PATMOS-x level-3) -5.9 % (MODIS) -3.4 % (ISCCP)
Cloud Top Height Cloud Top Pressure	(CTH) (CTP)	800 m 50 hPa	-840 m (CALIPSO level-2) -4.3 hPa (PATMOS-x level-2b) -18 hPa (PATMOS-x level-3) -88 hPa (MODIS) 13 hPa (ISCCP)
Cloud Phase	(CPH)	10 % (absolute)	1-2 % (PATMOS-x) 1-5 % (MODIS) 1-9 % (ISCCP)
Liquid Water Path	(LWP)	10 gm^{-2}	-3.5 to -0.5 gm^{-2} % (UWisc) 4.3 gm^{-2} (PATMOS-x) -2.3 to 3.3 gm^{-2} (MODIS) 10 to 17 gm^{-2} (ISCCP)
Ice Water Path	(IWP)	20 gm^{-2}	0.6 to 5.1 gm^{-2} (MODIS) 7.4 to 8.6 gm^{-2} (ISCCP)
Joint Cloud Histogram	(JCH)	n/a	n/a

-

Table 7.2 *Summary of validation results compared to target precisions for each cloud product. Consistency checks marked in blue.*

Product		Precision requirement (bc-RMS)	Achieved precisions
Cloud Fractional Cover	(CFC)	20 % (absolute)	40 %(CALIPSO level-2) 6.7 % (SYNOP level-3) 1.6 % (PATMOS-x level-2b/level-3) 10 % (PATMOS-x level-3) 7.3 % (MODIS) 8.8 % (ISCCP)
Cloud Top Height Cloud Top Pressure	(CTH) (CTP)	1700 m 100 hPa	2380 m (CALIPSO) 11 hPa (PATMOS-x level-2b/level-3) 85 hPa (PATMOS-x level-3) 58 hPa (MODIS) 93 hPa (ISCCP)
Cloud Phase	(CPH)	20 % (absolute)	6-7 % (PATMOS-x) 8-9 % (MODIS) 13-16 % (ISCCP)
Liquid Water Path	(LWP)	20 gm^{-2}	11-20 gm^{-2} (UWisc) 17 gm^{-2} (PATMOS-x) 8-11 gm^{-2} (MODIS) 14-19 gm^{-2} (ISCCP)
Ice Water Path	(IWP)	40 gm^{-2}	20-24 gm^{-2} (MODIS) 25-31 gm^{-2} (ISCCP)
Joint Cloud Histogram	(JCH)	n/a	n/a

Table 7.3 *Summary of validation results compared to target decadal stabilities for each cloud product. Consistency checks marked in blue.*

Product		Decadal stability requirement (change per decade)	Achieved stabilities
Cloud Fractional Cover	(CFC)	2 % (absolute)	-1.3 % (SYNOP) n/a (CALIPSO) 0.2 % (PATMOS-x) -1.1 % (MODIS)
Cloud Top Height	(CTH)	200 m	n/a (CALIPSO)
Cloud Top Pressure	(CTP)	20 hPa	-4.0 hPa (MODIS)
Cloud Phase	(CPH)	2 % (absolute)	1.6-2.2 % (MODIS)
Liquid Water Path	(LWP)	3 gm^{-2}	1.0-2.3 gm^{-2} (MODIS)
Ice Water Path	(IWP)	6 gm^{-2}	2.7-3.7 gm^{-2} (MODIS)
Joint Cloud Histogram	(JCH)	n/a	n/a

Final Remarks

There are already several satellite-based climate data records available providing similar information. However, in our opinion the added value of the CM SAF data record is:

- Cf. MODIS: much longer record (34 years vs 13 years)
- Cf. ISCCP: more homogeneous (no GEO used) and more spectral channels used
- Cf. PATMOS-x: good to have two similar data records produced with different algorithms to identify strengths /weaknesses of both approaches
- Cf. CALIPSO, SSM-I, UWisc: difference measurement principles, different variables measured, longer time frame
- The availability of additional surface radiation and surface albedo products produced from the same original data

Finally, it should be emphasised that the CLARA-A2 processing effort included not only significant algorithm improvements but also an unprecedented and rigorous (compared to CLARA-A1) quality control procedure of the original AVHRR GAC level-1b data record. In this respect the new data record appears to be much more stable and robust compared to CLARA-A1 and even compared to data records such as PATMOS-x. This is also a consequence of the in-depth nature of all validation efforts and the execution of the imposed feedback loop recommended at the DR1-5 review for CLARA-A1. This has led to some delays in the processing but it has enabled early discovery and correction of some crucial weaknesses of both technical and scientific nature.

8 References

Baum, B. A., W.P. Menzel, R.A. Frey, D.C. Tobin, R.E. Holz, S.A. Ackermann, A.K. Heidinger and P. Yang, 2012: MODIS Cloud-Top Property Refinements for Collection 6. *J. Appl. Meteor. Clim.*, **51**, 1145-1163., DOI: http://dx.doi.org/10.1175/JAMC-D-11-0203.1

Baran, Anthony. J., Shcherbakov, V. N., Baker, B. A., Gayet, J. F. and Lawson, R. P., 2005: On the scattering phase-function of non-symmetric ice-crystals. *Q.J.R. Meteorol. Soc.*, 131: 2609–2616. doi: 10.1256/qj.04.137.

Chepfer, H., S. Bony, D. Winker, G. Cesana, J. L. Dufresne, P. Minnis, C. J. Stubenrauch, and S. Zeng, 2010: The GCM Oriented CALIPSO CloudProduct (CALIPSO-GOCCP), *J. Geophys. Res.*, **115**, D00H16,doi:10.1029/2009JD012251.

Delanoë, J., and R. J. Hogan, 2008: A variational scheme for retrieving ice cloud properties from combined radar, lidar, and infrared radiometer, *J. Geophys. Res.*, **113**, D07204, doi:10.1029/2007JD009000.

Eliasson, S., G. Holl, S. A. Buehler, T. Kuhn, M. Stengel, F. Iturbide-Sanchez, *and* M. Johnston, 2013*:* Systematic and random errors between collocated satellite ice water path observations, *J. Geophys. Res. Atm.*, **118**, 2629–2642, *doi:*10.1029/2012JD018381.

Eliasson, S., S. A. Buehler, M. Milz, P. Eriksson, and V. O. John, 2011: Assessing observed and modelled spatial distributions of ice water path using satellite data, Atmos. Chem. Phys., 11, 375–391, doi:10.5194/acp-11-375-2011.

GCOS, 2006: Systematic observation requirements for satellite-based products for climate, http://www.wmo.int/pages/prog/gcos/Publications/gcos-107.pdf

González A., 2009: Measurement of Areas on a Sphere Using Fibonacci and Latitude--Longitude Lattices. Mathematical Geosciences. 42 (1), 49-64. doi:10.1007/s11004-009-9257-x

Hamann, U., et al., 2014: Remote sensing of cloud top pressure/height from SEVIRI: analysis of ten current retrieval algorithms, *Atm. Meas. Tech.*, **7**, 2839-2867, doi:10.5194/amt-7-2839-2014.

Karlsson, K.-G., 2003: A ten-year cloud climatology over Scandinavia derived from NOAA AVHRR imagery. *Int. J. Climatol.*, 23, 1023-1044.

Karlsson, K. -G., & Johansson, E. (2013). On the optimal method for evaluating cloud products from passive satellite imagery using CALIPSO–CALIOP data: example investigating the CM SAF CLARA-A1 dataset. *Atm. Meas. Tech.*, 6, 1271–1286, http://dx.doi.org/10.5194/amt-6-1271-2013.

Heidinger, A.K., W.C. Straka, C.C. Molling,J.T. Sullivan and X.Q. Wu, 2010:Deriving an inter-sensor consistent calibration for the AVHRR solar reflectance data record. *Int. J. Rem. Sens.*, 31, 6493-6517.

Heidinger, A.K., M.J. Pavolonis, 2009: Gazing at Cirrus Clouds for 25 Years through a Split Window. Part I: Methodology. *J. Appl. Meteor. Climatol.*, 48, 1100–1116. doi: 10.1175/2008JAMC1882.1

Heidinger, A.K., M.D. Goldberg, A. Jelenak and M.J. Pavolonis, 2005: A new AVHRR cloud climatology, *Proc. SPIE 5658*, 197, doi: 10.1117/12.579047.

Heidinger, A. K., Evan, A. T., Foster, M., and Walther, A., 2012: A Naïve Bayesian cloud detection scheme derived from CALIPSO and applied within PATMOS-x. Journal of Applied Meteorology and Climatology, 51, 1129–1144.

Heidinger, A.K., M.J. Foster, A. Walther and X. Zhao, 2014: The Pathfinder Atmospheres-Extended AVHRR Climate Dataset, *Bull. Amer. Met. Soc.*, June 2014, 909-922. Doi: 10.1175/BAMS-D-12-00246.1

Hess, H., R. B. A. Koelemeijer, and P. Stammes, 1998: Scattering matrices of imperfect hexagonal crystals. *J. Quant. Spectrosc. Ra*., 60, 301–308.

Ignatov, A., I. Laszlo, E.D. Harrod, K.B. Kidwell, and G.P. Goodrum, 2004: Equator crossing times for NOAA, ERS, and EOS sun-synchronous satellites. Int . J. Remote Sensing, 25, 5255–5266, doi:10.1080/01431160410001712981.

Mittaz, P.D. and R. Harris, 2009: A Physical Method for the Calibration of the AVHRR/3 Thermal IR Channels 1:The Prelaunch Calibration Data.*J. Atmos.Ocean. Tech.,* 26**,** 996-1019*,* doi: 10.1175/2008JTECHO636.1

Nakajima, T. and M.D. King, 1990: Determination of optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory, *Journal of the Atmospheric Sciences*, 47, pp. 1878–1893.

Pavolonis, M.J., A.K. Heidinger, T.Uttal, 2005: Daytime Global Cloud Typing from AVHRR and VIIRS: Algorithm Description, Validation, and Comparisons. *J. Appl. Meteor*., 44, 804–826. doi: 10.1175/JAM2236.1

Pincus R., S. Platnick, S. A. Ackerman, R. S. Hemler and R. J. P. Hofmann, 2012: Reconciling Simulated and Observed Views of Clouds: MODIS, ISCCP, and the Limits of Instrument Simulators. J. Climate, 25 (13), 4699-4720. doi:10.1175/JCLI-D-11-00267.1

Platnick, S., M. King, S. Ackerman, et al., 2003: The MODIS cloud products: Algorithms and examples from Terra., *IEEE Trans Geosci Remote Sens* 41 (2): 459-473.

O'Dell, C.W., F.J. Wentz, and R. Bennartz, 2008: Cloud Liquid Water Path from Satellite-Based Passive Microwave Observations: A New Climatology over the Global Oceans. J. Climate, 21, 1721–1739, doi:10.1175/2007JCLI1958.1

Rossow, W.B., and R.A. Schiffer, 1991: ISCCP cloud data products. Bull. Amer. Meteorol. Soc., 71, 2-20.

Rossow, W. B., and R. A. Schiffer, 1999: Advances in understanding clouds fromISCCP, Bull. Amer. Meeorot. Soc., 80, 2261-2287.

Rossow, W.B., A.W. Walker, D.E. Beuschel, and M.D. Roiter, 1996: International Satellite Cloud Climatology Project (ISCCP) Documentation of New Cloud Datasets. WMO/TD-No. 737, World Meteorological Organization.

Swinbank, R. and R. J. Purser, 2006: Fibonacci grids: A novel approach to global modelling. Quarterly Journal of the Royal Meteorological Society, 132 (619), 1769–1793. doi:10.1256/qj.05.227

Stubenrauch, C. J., W. B. Rossow, S. Kinne, S. Ackerman, G. Cesana, H. Chepfer, L. Di Girolamo, B. Getzewich, A. Guignard, A. Heidinger, B. Maddux, P. Menzel, P. Minnis, C. Pearl, S. Platnick, C. Poulsen, J. Riedi, S. Sun-Mack, A. Walther, D. Winker, S. Zeng, and G. Zhao, 2012: ASSESSMENT OF GLOBAL CLOUD DATASETS FROM SATELLITES: Project and Database initiated by the GEWEX Radiation Panel, Bull. Amer. Meteor. Soc., doi: 10.1175/BAMS-D-12-00117

Thomas, S.M., A.K. Heidinger, and M.J. Pavolonis, 2004: Comparison of NOAA's Operational AVHRR-Derived Cloud Amount to Other Satellite-Derived Cloud Climatologies. *J. Climate*, **17**, 4805–4822. doi: 10.1175/JCLI-3242.1

Vaughan, M., Powell, K., Kuehn, Rl.,Young, S., Winker, D., Hostetler, C., Hunt, W., Liu, Z., McGill, M., and Getzewich, B., 2009: Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements, J. Atmos. Oceanic Technol., 26, 2034–2050, doi: 10.1175/2009JTECHA1228.1.

Winker, D. M., Vaughan, M.A., Omar, A., Hu, Y, Powell, K.A., Liu, Z., Hunt, W.H., and Young, S.A., 2009: Overview of the CALIPSO mission and CALIOP data processing algorithms, J. Atmos. Oceanic. Technol., 26, 2310-2323, doi:10.1175/2009JTECHA1281.1.

World Meteorological Organization, 1989: Guide to Meteorological Instruments and Methods of Observation. WMO - No.8, Geneva.

9 Glossary

AMSR-E	Advanced Microwave Scanning Radiometer for EOS
ATBD	Algorithm Theoretical Baseline Document
AVHRR	Advanced Very High Resolution Radiometer
BC-RMS	Bias-Corrected RMS
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarisation
CDOP	Continuous Development and Operations Phase
CFC	Fractional Cloud Cover
CLARA-A	CM SAF cLoud, Albedo and Radiation products, AVHRR-based
CLAAS	CM SAF cLoud dAtAset using SEVIRI
CM SAF	Satellite Application Facility on Climate Monitoring
COT	Cloud Optical Thickness
CPH	Cloud Phase
CPR	Cloud Profiling Radar
CTH	Cloud Top Height
CTO	Cloud Top product
CTP	Cloud Top Pressure
CTT	Cloud Top Temperature
CPP	Cloud Physical Properties
DAK	Doubling Adding KNMI (radiative transfer model)
DRR	Delivery Readiness Review
DWD	Deutscher Wetterdienst (German MetService)
ECMWF	European Centre for Medium Range Forecast
ECV	Essential Climate Variable
ERA-Interim	Second ECMWF Re-Analysis dataset
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FAR	False Alarm Ratio
FCDR	Fundamental Climate Data Record
FCI	Flexible Combined Imager
GAC	Global Area Coverage (AVHRR)
GCOS	Global Climate Observing System

GSICS	Global Space-Based Inter-Calibration System
ISCCP	International Satellite Cloud Climatology Project
ITCZ	Inter Tropical Convergence Zone
IWP	Ice Water Path
JCH	Joint Cloud properties Histogram
KNMI	Koninklijk Nederlands Meteorologisch Instituut
KSS	Hanssen-Kuiper Skill Score
LWP	Liquid Water Path
MODIS	Moderate Resolution Imaging Spectroradiometer
MSG	Meteosat Second Generation
MTG	Meteosat Third Generation
NOAA	National Oceanic & Atmospheric Administration
NWC SAF	SAF on Nowcasting and Very Short Range Forecasting
NWP	Numerical Weather Prediction
PATMOS-x	Pathfinder Atmospheres-Extended dataset (NOAA)
POD	Probability Of Detection
PPS	Polar Platform System (NWC SAF polar cloud software package)
PRD	Product Requirement Document
PUM	Product User Manual
REFF	Cloud particle effective radius
RMS	Root Mean Square (Error)
RTTOV	Radiative Transfer model for TOVS
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SAF	Satellite Application Facility
SMHI	Swedish Meteorological and Hydrological Institute
SYNOP	Synoptic observations
SZA	Solar Zenith Angle
UWisc	University of Wisconsin passive microwave based LWP data record
VZA	Viewing Zenith Angle