

Validation Report
SEVIRI cloud products
CLAAS Edition 3

DOI: [10.5676/EUM_SAF_CM/CLAAS/V003](https://doi.org/10.5676/EUM_SAF_CM/CLAAS/V003)

	TCDR	ICDR
Fractional Cloud Cover (CFC)	CM-21014	CM-5011
Joint Cloud property Histogram (JCH)	CM-21023	CM-5021
Cloud Top level (CTO)	CM-21033	CM-5031
Cloud Phase (CPH)	CM-21043	CM-5041
Liquid Water Path (LWP)	CM-21053	CM-5051
Ice Water Path (IWP)	CM-21063	CM-5061

Reference Number:
Issue/Revision Index:
Date:

SAF/CM/KNMI/VAL/SEV/CLD
3.1
08.08.2022

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

Document Signature Table

	Name	Function	Signature	Date
Authors	Jan Fokke Meirink Martin Stengel Nikos Benas Irina Solodovnik Nina Håkansson Karl-Göran Karlsson	CM SAF scientist (KNMI) CM SAF scientist (DWD) CM SAF scientist (KNMI) CM SAF scientist (DWD) CM SAF scientist (SMHI) CM SAF scientist (SMHI)		08/08/2022
Editor	Marc Schröder	Science Coordinator		08/08/2022
Approval	CM SAF Steering Group			xx/xx/2022
Release	Rainer Hollmann	Project Manager		xx/xx/2022

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
3.0	23/05/2022	SAF/CM/KNMI/VAL/SEV/CLD/3.0	First official version

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD
		Issue: 3.1
		Date: 08.08.2022

			submitted for DRR 3.7
3.1	08/08/2022	SAF/CM/KNMI/VAL/SEV/CLD/3.1	Version after DR/ORR, with RID responses implemented

Applicable Documents

Reference	Title	Code / Date
AD 1	EUMETSAT CM SAF CDOP 3 Product Requirements Document (PRD)	SAF/CM/DWD/PRD/4.0
AD 2	Requirements Review SEVIRI Edition 3 data records (CLAAS Edition 3)	SAF/CM/CDOP3/KNMI/RR37

Reference Documents

Reference	Title	Code
RD 1	Product User Manual SEVIRI Cloud Products Edition 3 (CLAAS-3)	SAF/CM/DWD/PUM/SEV/CLD/3.0
RD 2	Algorithm Theoretical Basis Document SEVIRI cloud products processing chain	SAF/CM/DWD/ATBD/SEV/CLD/3.2
RD 3	Algorithm Theoretical Basis Document (An Appendix to the NWC/PPS) Cloud Probability and Cloud Top Temperature/Height from SEVIRI	SAF/CM/SMHI/ATBD/SEV/PPSSEV/3.2
RD 4	Algorithm Theoretical Basis Document SEVIRI cloud physical products	SAF/CM/KNMI/ATBD/SEV/PPP/3.2
RD 5	Validation Report SEVIRI cloud products CLAAS Edition 2	SAF/CM/DWD/VAL/SEV/CLD/2 v1.1, dated 10/06/2016.

Table of Contents

1	Executive Summary	15
2	The EUMETSAT SAF on Climate Monitoring	19
3	Introduction to CLAAS-3.....	21
4	Validation strategy.....	24
5	Brief characterization of systematic changes from CLAAS-2.1 to CLAAS-3.....	27
5.1	Time series of L3 products.....	27
5.2	Spatial maps of L3 products.....	31
5.3	Some characteristics of L2 products and their diurnal cycle	34
5.4	Summary of differences between CLAAS-2.1 and CLAAS-3	38
6	Evaluation of CLAAS-3 instantaneous (L2) cloud parameters	39
6.1	Validation with CALIOP.....	39
6.1.1	Cloud Mask and Cloud Phase	40
6.1.2	Cloud Top Level	45
6.1.3	Summary of validation with CALIOP	50
6.2	Validation with DARDAR.....	50
6.3	Validation with AMSR2	53
6.4	Comparison with MODIS	57
6.4.1	Cloud top height	57
6.4.2	Cloud optical thickness	59
6.4.3	Cloud particle effective radius	61
6.4.4	Liquid cloud droplet number concentration and geometrical thickness.....	64
6.4.5	Summary of comparison with MODIS	65
7	Evaluation of CLAAS-3 aggregated (L3) cloud parameters	67
7.1	Validation with SYNOP	67
7.2	Validation with CALIOP.....	71
7.2.1	Cloud Fractional Cover (CFC)	72
7.2.2	Cloud Top Height (CTH)	75

7.2.3	Summary of results.....	76
7.3	Validation with MAC-LWP	77
7.4	Comparison with MODIS	81
7.4.1	Cloud Fractional Cover (CFC)	82
7.4.2	Cloud Top Level	83
7.4.3	Cloud Thermodynamic Phase (CPH).....	85
7.4.4	Liquid Water Path (LWP)	86
7.4.4.1	Liquid COT and CRE	87
7.4.4.2	Liquid CDNC.....	90
7.4.5	Ice Water Path (IWP).....	91
7.4.5.1	Ice COT and CRE	92
7.4.6	JCH: Joint Cloud property Histograms of CTP and COT.....	94
7.4.7	Summary of results.....	95
8	Evaluation of CLAAS-3 ICDR	97
8.1	L2	97
8.2	L3	100
9	Conclusions.....	105
10	Appendix A: SEVIRI L1.5 data.....	108
11	Appendix B: Reference datasets	111
11.1	SYNOP: manual cloud observations from surface stations	111
11.2	CALIPSO-CALIOP	112
11.3	DARDAR	115
11.4	MODIS.....	116
11.5	Microwave imagers.....	116
12	References.....	118
13	Acronyms.....	121

List of Tables

Table 1-1: Summary of CLAAS-3 validation results compared to requirements for each cloud product. Required and achieved accuracies are formulated in terms of bias, precisions in terms of bc-rmsd (except for CFC and CPH L2, for which the metric is the Hanssen-Kuipers Skill Score, KSS), and stabilities in terms of decadal trend in bias. All numbers, except KSS, are in the units indicated for the respective cloud products. Validation results are color-coded as follows: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirement. CALIOP results are reported for different values of (I)COT. Evaluations against MODIS are indicated in grey since they are viewed as consistency checks rather than true validation. CPH, LWP and IWP validation scores are for the 3.9- μ m-based products.	17
Table 3-1: CLAAS-3 TCDR and ICDR products with formal product IDs and abbreviations according to [AD 1].	21
Table 4-1: Contingency table for the 2x2 problem. <i>n_{ij}</i> is the number of cases where CLAAS-3 reports event i and the reference reports event j. For example event 1 may be clear and event 2 may be cloudy.	24
Table 4-2: CM ^o SAF CLAAS-3 products and their respective target requirements for accuracy, precision and stability (defined in [AD 1]). L2 requirements refer to the pixel level, L3 requirements to daily/monthly means. The requirements are applicable to the SEVIRI disk.	25
Table 5-1: Percentage of liquid and ice CRE retrievals below and above the look-up table, averaged over the full disk on 25 March 2013.	37
Table 6-1: Overview of reference datasets used for the evaluation of CLAAS-3 L2 parameters.	39
Table 6-2: Summary of Hanssen-Kuipers Skill Scores for CLAAS-3 cloud mask and cloud phase. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.	44
Table 6-3: Overall validation results for CLAAS-3 cloud top products. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.	49
Table 6-4: Summary of CLAAS IWP validation scores with reference to DARDAR v3.00 and v3.10. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.	53
Table 6-5: Summary of CLAAS LWP validation scores with reference to AMSR2. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.	56
Table 6-6: Summary of CLAAS-MODIS comparison scores: mean difference, bc-rmsd, and correlation coefficient. Where applicable, results for the mean difference and bc-rmsd are color-coded with reference to the L2 requirements: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.	66
Table 7-1: Overview of reference datasets used for the evaluation of CLAAS-3 L3 parameters.	67
Table 7-2: Time series averages of bias and bc-rmse, and stability for CLAAS-3 CFC that were evaluated based on corresponding SYNOP monthly mean data. Remark: The anticipated error of SYNOP observations is probably of the order of 10 %, i.e., close to the Threshold requirement. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.	71


	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

Table 7-3: Time series averages of bias and bc-rmsd, and stability for total, low-, mid- and high-level CFC and CTH from CLAAS-3 data record that were evaluated based on corresponding CALIPSO-GEWEX L3 data. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements. 77

Table 7-4: Overall requirement compliance of the CLAAS-3 all-sky LWP product with respect to the bias, bc-rmsd, and stability of the bias, calculated over the marine stratocumulus region (10°-20° S, 0°-10°E) for validation with MAC-LWP data. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements. 80

Table 7-5: Time series averages of bias and bc-rmsd, and stability for all CLAAS-3 data records that were evaluated based on corresponding MODIS L3 data. Scores are color-coded where applicable: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements. 95

Table 8-1: Bias and bias-corrected root mean square error for L2 cloud top pressure and cloud water path; hit rate and Hansen-Kuipers skill score for L2 cloud mask and phase for the time slots shown in figures above. Scores are color-coded where applicable: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal (TCDR) requirements. 100

Table 8-2: Evaluation scores of CLAAS-3 TCDR and CLAAS-3 ICDR comparison. Scores are color-coded where applicable: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal (TCDR) requirements. 104

List of Figures

Figure 3-1: Illustration of the CLAAS-3 TCDR and ICDR temporal coverage.....	21
Figure 3-2: Overview of the SEVIRI measurement record used as basis for the generation of the CLAAS-3 TCDR. The top panel shows which satellites were used with short-term data gaps enlarged by a factor 5 for better visibility. The bottom panel shows the location of the satellites along the equator.	22
Figure 3-3: SEVIRI spatial coverage and on-ground resolution expressed as the edge length of a square having the same area as the SEVIRI grid cell.....	23
Figure 5-1: Time series of the CLAAS-2.1 and CLAAS-3 CFC products (left column) and time series of the difference between both data sets (right column).	28
Figure 5-2: Time series of the CLAAS-2.1 and CLAAS-3 CPH products (left column) and time series of the difference between both data sets (right column).	28
Figure 5-3: Time series of the CLAAS-2.1 and CLAAS-3 CTO products (left column) and time series of the difference between both data sets (right column).	29
Figure 5-4: Time series of the CLAAS-2.1 and CLAAS-3 LWP products (left column) and time series of the difference between both data sets (right column).	30
Figure 5-5: Time series of the CLAAS-2.1 and CLAAS-3 IWP products (left column) and time series of the difference between both data sets (right column).	31
Figure 5-6: Maps of averaged cfc, cph, ctp, lwp_allsky, and iwp_allsky from CLAAS-2.1 (left column), CLAAS-3 (middle column) and the corresponding differences (right column) for the overlapping period: 01/02/2004 – 31/12/2017.....	33
Figure 5-7: Diurnal cycle of cloud CLAAS-2.1 and CLAAS-3 cloud properties averaged over a 100x100 pixel region in the southeast Atlantic on 25 March 2013. Twilight conditions (SZA between 75 and 95 degrees) are indicated by the shaded area. The middle vertical dotted line represents cloud glory conditions (scattering angle near 180 degrees), while the other two vertical dotted lines represent cloud bow conditions (scattering angle near 140 degrees).	35
Figure 5-8: Diurnal cycle of the fraction of COT-CRE retrievals below the look-up table (top) and above the look-up table (bottom), for liquid clouds (left) and ice clouds (right) on 25 March 2013. The full disk is considered with the requirement that 70% of the pixels satisfies the retrieval criteria for solar and satellite zenith angles.	36
Figure 5-9: Histograms of cloud top height, cloud optical thickness, liquid cloud droplet effective radius, and ice cloud particle effective radius from CLAAS-2.1 and CLAAS-3 full-disk retrievals on 25 March 2013.....	37
Figure 6-1: Example matchup of CALIOP track and SEVIRI scan. The labels display the nominal timestamp of the SEVIRI scan that was matched to the identically coloured ground track segment. ...	39
Figure 6-2: 10-day moving average of cloud mask (yielding a cloud fraction, top) and cloud phase (bottom, yielding an ice cloud fraction) from CLAAS-3 and CALIOP, averaged over the full disk. Dashed lines show COT-filtered CALIOP data.....	40
Figure 6-3: Left: CLAAS-3 cloud scores as a function of the COT threshold used to discriminate clear and cloudy CALIOP observations. KSS denotes the Hanssen-Kuipers Skill Score. Right: CLAAS-3	

phase scores as a function of the ICOT threshold, which determines the reference CALIOP cloud layer.
..... 41

Figure 6-4: Spatial distribution of total (upper row) and liquid (bottom row) cloud fraction as seen by CLAAS-3 (left) and CALIOP (middle), and their difference (right). CALIOP cloud criterion is total column COT > 0. CALIOP liquid phase is taken from the layer where ICOT exceeds 0.2..... 42

Figure 6-5: Spatial distribution of validation scores of CLAAS-3 against CALIOP: probability of detection (POD), false alarm ratio (FAR) and Hanssen-Kuipers Skill Score (KSS) for total cloud fraction (upper panel) and liquid cloud fraction (bottom panel). CALIOP cloud fraction is taken for COT > 0, and liquid phase is taken from the layer where ICOT exceeds 0.2. 43

Figure 6-6: CLAAS-3 cloud mask and phase Hanssen-Kuipers Skill score (KSS) in 4-degree latitude bands. Darker and lighter red shadings indicate the target and threshold KSS requirements, respectively. Top: Solid/dashed lines were calculated using total column COT > 0/0.2 as cloud criterion. Bottom: Solid lines refer to the uppermost CALIOP cloud layer, dashed lines refer to the CALIOP cloud layer where ICOT exceeds the 0.2 threshold. 44

Figure 6-7: Ten-day moving average of CLAAS-3 and CALIOP cloud top products averaged over the full disk. Solid black line denotes the uppermost CALIOP cloud layer; dashed black line represents the CALIOP cloud layer where ICOT exceeds the 0.2 threshold. 45

Figure 6-8: Mean cloud top bias (CLAAS-3 – CALIOP) compared against the CALIOP cloud layer where ICOT exceeds a certain threshold. Error bars represent the bias-corrected root mean square error. . 46

Figure 6-9: Zonal averages of CLAAS-3 and CALIOP cloud top products in 4-degree latitude bands. Solid black line denotes the uppermost CALIOP cloud layer; dashed black line represents the CALIOP cloud layer where ICOT exceeds the 0.2 threshold. 46

Figure 6-10: 1D-histograms of cloud top products from CLAAS-3 and CALIOP. Solid lines denote liquid phase, dotted lines represent ice clouds. 47

Figure 6-11: Correlation of cloud top products between CLAAS-3 and CALIOP. A dotted line marks the diagonal; red lines show the result of a least squares linear fit. The textbox in the upper left corner displays the Pearson correlation coefficient, the text in the lower right corner the linear fit function. CALIOP values are taken from the layer where ICOT exceeds 0.2..... 48

Figure 6-12: Spatial distribution of cloud top products from CLAAS-3 and CALIOP, and their difference. CALIOP values are taken from the layer where ICOT exceeds 0.2..... 49

Figure 6-13: Comparison between CLAAS and DARDAR v3.00 ice cloud optical thickness: scatter density plots of CLAAS-3 1.6- μm (left) and CLAAS-3 3.9- μm (centre) ice COT versus DARDAR, and 1D-histograms of CLAAS-3, CLAAS-2.1 and DARDAR ice COT (right). Statistics of the inter-comparison are included in the plots: n is the number of collocations, μ is the mean, σ is the standard deviation or bc-rmsd, and r is the linear Pearson correlation coefficient. In the scatter density plots linear (in log-space) orthogonal fit lines are also included. Note that the two CLAAS-3 COT flavors are nearly identical and their histograms cannot be distinguished. See text for details regarding the collocation procedure.
..... 51

Figure 6-14: As Figure 6-13 but for the ice cloud particle effective radius. For DARDAR a CRE weighted near the cloud top ($\tau_w = 1$) was derived, as explained in the text..... 52

Figure 6-15: As Figure 6-13 but for ice water path..... 52

Figure 6-16: Comparison between CLAAS and AMSR2 liquid water path: scatter density plots of CLAAS-3 1.6- μm (top left), CLAAS-3 3.9- μm (top centre) and CLAAS-2.1 1.6- μm (top right) based LWP versus AMSR2, and spatial distribution of CLAAS-3, AMSR2, and their difference for 1.6 μm (middle row) and 3.9 μm (bottom row). Statistics of the inter-comparison are included in the plots: n is the number of collocations, μ is the mean, σ is the standard deviation or bc-rmsd, and r is the linear Pearson correlation coefficient. In the scatter density plots linear orthogonal fit lines are also included. See text for details regarding the collocation procedure. 54

Figure 6-17: Same as in Figure 6-16 but restricted to non-raining pixels. 55

Figure 6-18: Scatter density plots of CLAAS-3 1.6- μm (left), CLAAS-3 3.9- μm (centre) and CLAAS-2.1 1.6- μm (right) based LWP versus AMSR2 for non-raining and fully cloudy pixels. 56

Figure 6-19: CLAAS LWP error estimates versus the absolute difference between CLAAS and AMSR2 LWP: CLAAS-3 1.6 μm (left), CLAAS-3 3.9 μm (centre), and CLAAS-2.1 1.6 μm (right). The collocation dataset is the same as in Figure 6-16. 56

Figure 6-20: Comparison between CLAAS and MODIS cloud top height of liquid clouds: scatter density plot of CLAAS-3 versus MODIS C6.1 (top left), 1D-histograms of CLAAS-3, CLAAS-2.1 and MODIS (top right), and spatial distribution of CLAAS-3, MODIS, and their difference (bottom left to right). Statistics of the inter-comparison are included in the plots: n is the number of collocations, μ is the mean, σ is the standard deviation, and r is the linear Pearson correlation coefficient. In the scatter density plot a linear orthogonal fit line is also included. See text for details regarding the collocation procedure..... 58

Figure 6-21: As Figure 6-20 but for ice clouds. 59

Figure 6-22: As Figure 6-20 but for liquid cloud optical thickness. The scatter density plot (top left) and maps (bottom) are compiled from the 3.9- μm based CLAAS-3 COT and the 3.7- μm based MODIS logarithmically-averaged COT. 1D-histograms (top right) are shown for retrievals based on different SWIR channels as indicated, and for linearly and logarithmically (indicated by 'LOG') averaged MODIS retrievals. 60

Figure 6-23: Liquid cloud optical thickness as a function of solar zenith angle: for CLAAS and MODIS separately (left) and CLAAS-MODIS difference (right). The normalized number of observations per solar zenith angle bin is indicated by the red line in both panels. 60

Figure 6-24: As Figure 6-22 but for ice cloud optical thickness. 61

Figure 6-25: Comparison between CLAAS and MODIS liquid cloud droplet effective radius: scatter density plot of CLAAS-3 versus MODIS C6.1 CRE based on 1.6 μm (top left), and on 3.9/3.7 μm (top middle), 1D-histograms of CLAAS-3, CLAAS-2.1 and MODIS CRE (top right), spatial distribution of CLAAS-3, MODIS, and their difference for CRE based on 1.6 μm (middle row), and on 3.9/3.7 μm (bottom row). 62

Figure 6-26: Liquid cloud droplet effective radius as a function of solar zenith angle: for CLAAS and MODIS separately (left) and CLAAS-MODIS difference (right). The normalized number of observations per solar zenith angle bin is indicated by the red line in both panels. 63

Figure 6-27: As Figure 6-25 but for the ice cloud particle effective radius. 64

Figure 6-28: Comparison between CLAAS and MODIS liquid cloud droplet number concentration: scatter density plot of CLAAS-3 versus MODIS C6.1 CDNC (top left), 1D-histograms of CLAAS-3 and

MODIS CDNC (top right), spatial distribution of CDNC from CLAAS-3, MODIS, and their difference (bottom row). 65

Figure 6-29: As Figure 6-28, but for liquid cloud geometrical thickness. 65

Figure 7-1: Mean difference between CLAAS-3 and SYNOP cloud cover at each preselected SYNOP site for the entire period 2004-2020. 68

Figure 7-2: Time series of mean cloud cover for CLAAS-3 (red), 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 2004-2020. 69

Figure 7-3: Validation of monthly mean CFC from CLAAS-3 with SYNOP: scatter plot (left) and frequency distribution of deviation (right). 70

Figure 7-4: Time series of the bias between the CLAAS-3 and the SYNOP cloud fractional cover monthly mean. The red line is the linear fit. 70

Figure 7-5: Maps of averaged cloud fractional cover based on the overlapping time period of CLAAS-3 and CALIOP L3 (2006/06 – 2016/12) from CLAAS-3 (left column), top layer flavor of CALIPSO-GEWEX cloud product (middle column) and their difference (right column). 72

Figure 7-6: Averaged zonal mean CFC (left), and time series of CFC and bias (right) from CLAAS-3 and the CALIPSO-GEWEX top layer and passive flavor products. 72

Figure 7-7: Maps of averaged fractional cover of low (top), middle (middle), and high (bottom) clouds based on the time period 2006/06 – 2016/12 from CLAAS-3 (left column), top layer flavor of CALIPSO-GEWEX cloud product (middle column) and their difference (right column). 74

Figure 7-8: Averaged zonal mean CFC (left), and time series of CFC and bias (right) from CLAAS-3 and CALIPSO-GEWEX top layer and passive flavor for low-level clouds (upper panel), mid-level clouds (middle panel) and high clouds (bottom panel). 75

Figure 7-9: Maps of averaged cloud top height based on the time period 2006/06 – 2016/12 from CLAAS-3 (left column), top layer flavor of CALIPSO-GEWEX cloud product (middle column) and their difference (right column). 76

Figure 7-10: Averaged zonal mean CFC (left) and time series of CTH and bias (right) from CLAAS-3 and CALIPSO-GEWEX top layer and passive flavor. 76

Figure 7-11: The marine stratocumulus region off the Namibian coast (0-10°E, 10-20°S), selected for the validation of CLAAS-3 all-sky LWP monthly mean diurnal cycle against corresponding MAC-LWP data. 78

Figure 7-12: Time series of the spatially averaged all-sky LWP at the marine Sc region (0-10°S, 10-20°E) from CLAAS-3 and MAC-LWP data (top plot). The values are averages of the 1-hour time slots between 7:00 UTC and 16:00 UTC. The corresponding bias and bc-rmsd are shown in the bottom plot, along with the bias trend during the period examined. 79

Figure 7-13: Monthly mean diurnal average of all-sky LWP from CLAAS-3 and MAC-LWP (top plot), spatially averaged at the marine Sc region (0-10°S, 10-20°E), along with their bias and bc-rmsd (bottom plot). 80

Figure 7-14: Maps of average CFC (top row), CFC_Day (middle row) and CFC_Night (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12) from CLAAS-3 (left column), MODIS (middle column) and their difference (right column). 83

Figure 7-15: Time series of CFC (top row), CFC_Day (middle row) and CFC_Night (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (% per decade) based on linear regression. 83

Figure 7-16: Maps of average CTH (top row) and CTP (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12) from CLAAS-3 (left column), MODIS (middle column) and their difference (right column). 84

Figure 7-17: Time series of CTH (top row) and CTP (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (m and hPa per decade) based on linear regression. 84

Figure 7-18: Maps of average cloud phase (CPH, top row) and day-only cloud phase (CPH_Day, bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12) from CLAAS-3 (left column), MODIS (middle column) and their difference (right column). 85

Figure 7-19: Time series of CPH (top row) and CPH_Day (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (% per decade) based on linear regression. 86

Figure 7-20: Maps of average all-sky Liquid Water Path (LWP) retrieved based on the 3.9 μm channel (top row) and on the 1.6 μm channel (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left column), MODIS (middle column) and their difference (right column). 87

Figure 7-21: Time series of all-sky LWP retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (g m⁻² per decade) based on linear regression. 87

Figure 7-22: Maps of average liquid Cloud Optical Thickness (COT) retrieved based on the 3.9 μm channel, and on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left column), MODIS (middle column) and their difference (right column). 88

Figure 7-23: Maps of average liquid Cloud Effective Radius (CRE) retrieved based on the 3.9 μm channel (top row) and on the 1.6 μm channel (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left column), MODIS (middle column) and their difference (right column). 88

Figure 7-24: Time series of liquid COT retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (per decade) based on linear regression. 89

Figure 7-25: Time series of liquid CRE retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (μm per decade) based on linear regression. 89

Figure 7-26: Maps of average liquid Cloud Droplet Number Concentration (CDNC) based on the period 2004/02 – 2015/12, from CLAAS-3 (left column), MODIS (middle column) and their difference (right column). In the top row, all available data are used, with the requirement that the time series is at least 50% complete on a grid cell basis. In the bottom row, only grid cells where both CLAAS-3 and MODIS are available were used in the averaging. 90

Figure 7-27: Time series of liquid Cloud Droplet Number Concentration (CDNC), retrieved based on the 3.9 μm channel, from monthly L3 CLAAS-3 and MODIS data (left plot) and corresponding bias and bc-rmsd (right plot). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E, and over sea only, due to lack of MODIS data over land. The blue dotted line in the bias plot shows the change in bias (cm^{-3} per decade) based on linear regression. 91

Figure 7-28: Maps of average all-sky Ice Water Path (IWP) retrieved based on the 3.9 μm channel (top row) and on the 1.6 μm channel (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left column), MODIS (middle column) and their difference (right column). . 91

Figure 7-29: Time series of all-sky IWP retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (g m^{-2} per decade) based on linear regression. 92

Figure 7-30: Maps of average ice Cloud Optical Thickness (COT) retrieved based on the 3.9 μm channel and based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left), MODIS (column) and their difference (right). 92

Figure 7-31: Maps of average ice Cloud Effective Radius (CRE) retrieved based on the 3.9 μm channel (top row) and on the 1.6 μm channel (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left column), MODIS (middle column) and their difference (right column). 93

Figure 7-32: Time series of ice COT retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (per decade) based on linear regression. 93

Figure 7-33: Time series of ice CRE retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (μm per decade) based on linear regression. 94

Figure 7-34: Joint Cloud Histograms (JCH) of CTP and COT from CLAAS-3 (left) and MODIS (middle, sum of Terra and Aqua counts), using the ISCCP heritage histogram bins, based on the entire CLAAS-3 time series (2004/01 – 2020/12). The rightmost JCH is also from MODIS, but includes partly cloudy (PCL) retrievals. All JCHs are normalized to 100%. 95

Figure 8-1: Maps of CLAAS-3 ICDR (left), CLAAS-3 TCDR (middle) and differences (right) of L2 variables on 10.01.2020 at 12 UTC..... 98

Figure 8-2: Maps of CLAAS-3 ICDR (left), CLAAS-3 TCDR (middle) and differences (right) of L2 variables on 10.07.2020 at 12 UTC..... 99

Figure 8-3: Maps of CLAAS-3 ICDR (left), CLAAS-3 TCDR (middle) and differences (right) of L3 monthly averaged products in January 2020. 101

Figure 8-4: As Figure 8-3 but for July 2020..... 102

Figure 8-5: Time series and bias-corrected rmse (left), zonal mean with global averaged scores (right) of CLAAS-3 ICDR and CLAAS-3 TCDR L3 monthly averaged products for 01/01/2020-31/12/2020. 104

Figure 10-1: Time series of brightness temperature measured at 00:00 UTC in SEVIRI IR channels. The measurements have been averaged over a 2112 x 2112 pixel area centred around 0°/0° latitude/longitude, which corresponds to about (30 °W – 30 °E , 30 °S – 30 °N). The solid lines show 1-year running mean BT, consecutively offset by 1 K, as indicated by the corresponding dotted lines. Vertical, dashed grey lines indicate the transitions between the prime MSG satellites. 108

Figure 10-2: Difference between SEVIRI IR channels: BT 3.9 – BT 10.8 μm (top), BT 6.2 – BT 10.8 μm (top), BT 12.0 – BT 10.8 μm (top). Blue lines show weekly values, while the orange lines are 1-year running means. Data have been spatially averaged as outlined in Figure 10-1. Vertical, dashed grey lines indicate the transitions between the prime MSG satellites. 109

Figure 11-1: The Aqua-Train satellites. (Image credit: NASA)..... 113

1 Executive Summary

This CM°SAF report provides information on the validation of the third edition of the CCloud property dAtaset using SEVIRI (CLAAS-3). All CLAAS-3 products are derived from measurements taken by the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) onboard the EUMETSAT Meteosat Second Generation (MSG) satellites. Four MSG satellites have been used: MSG-1 to MSG-4, also referred to as Meteosat-8 to Meteosat-11. CLAAS-3 consists of a Thematic Climate Data Record (TCDR), spanning the period from 2004 to 2020, and an Interim Climate Data Record (ICDR), starting in 2021 and extended operationally with low latency in the present. The ICDR is produced with the same algorithms as the TCDR but with a few differences in input data.

CLAAS-3 consists of the following cloud products:

	TCDR	ICDR
Fractional Cloud Cover (CFC)	CM-21014	CM-5011
Joint Cloud property Histogram (JCH)	CM-21023	CM-5021
Cloud Top level (CTO)	CM-21033	CM-5031
Cloud Phase (CPH)	CM-21043	CM-5041
Liquid Water Path (LWP)	CM-21053	CM-5051
Ice Water Path (IWP)	CM-21063	CM-5061

The reference datasets, used to evaluate the CLAAS-3 cloud products, were taken from ground-based observation sources (SYNOP) as well as from other satellite-based sources (e.g., CALIPSO, AMSR2, and MODIS). The comparisons with MODIS products have to be viewed as consistency checks because the MODIS type of measurements (passive visible and infrared imagery) are similar to SEVIRI. The comparisons to the other products come much closer to validation of the SEVIRI products because of their assumed superior measurement principle and quality, although these other products also have their uncertainties and pitfalls, which are outlined in this report.

The evaluation includes both Level 2 (L2) and Level 3 (L3) comparisons to not only demonstrate the applicability concerning the aggregated products, but also to allow a more in-depth analysis of the cloud properties provided on a high temporal resolution. Apart from the main products listed above also additional product layers, such as cloud optical thickness (COT), particle effective radius (CRE) and liquid cloud droplet number concentration (CDNC), are analysed.

All evaluations of the cloud products are done in the light of the product requirements. The evaluation scores and their compliance to the requirements for accuracy, precision, and stability are given in Table 1-1, and are briefly summarized in the following:

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

- **Fractional Cloud Cover (CFC)**

The CLAAS-3 L2 CFC product achieves the L2 target precision requirements compared to CALIOP. Validation with SYNOP observations shows that L3 CFC meets all target requirements, and even optimal requirements for accuracy and stability. Evaluation against CALIOP L3 shows that CLAAS-3 agrees best with the CALIOP ‘top flavor’ CFC product, meeting the target requirements except for precision. Finally, L3 comparisons with MODIS demonstrate very stable results (within optimal stability requirements), but a negative mean difference outside the target range.

- **Cloud Top level (CTO)**

The CLAAS-3 CTH product yields lower mean values (slightly worse than target requirement) than the ‘cloud top’ flavor of CALIOP but agrees well with the ‘passive flavor’, for which target bias requirements are fulfilled. Comparisons of mean L2 CTH with MODIS are within the optimal requirements for liquid clouds but worse than the target for ice clouds. CLAAS-3 ice cloud CTH is closer to CALIOP than MODIS. The stability of the CTH product meets the optimal requirements in comparison to both CALIOP and MODIS. CLAAS-3 CTP evaluations are overall consistent with those for CTH.

- **Cloud Thermodynamic Phase (CPH)**

The CLAAS-3 CPH product achieves the L2 target precision requirements compared to CALIOP. Comparisons of L3 CPH with MODIS fulfil the threshold requirements. A positive trend in the monthly differences is present but appears to be mainly caused by a negative trend in MODIS liquid cloud fraction.

- **Liquid Water Path (LWP)**

The CLAAS-3 LWP product meets L2 target requirements for bias and bc-rmsd against AMSR2, while the L2 comparisons with MODIS fulfil the threshold requirements. The L3 validation against LWP from microwave radiometers focusing on the south-east Atlantic shows that the target requirements are fulfilled. In comparison with MODIS, LWP satisfies the target requirements for precision and optimal requirements for accuracy and stability.

- **Ice Water Path (IWP)**

The CLAAS-3 IWP L2 product does not fulfil the threshold requirements against DARDAR observations, and only threshold requirements in comparison to MODIS. This is partly a consequence of the large range of IWP values, where differences between products can become large for the higher values, leading to relatively large bias and bc-rmsd even if correlations are reasonable. On the other hand, comparisons of L3 IWP with MODIS fulfil the optimal requirements for accuracy, precision and stability.

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

This product is excluded from specific requirement testing because it is composed of three already evaluated products (CPH, COT and CTP). Nevertheless, inter-

comparisons with the corresponding MODIS product show reasonable agreement. JCH provides added value to the underlying products by giving important clues on the statistical distribution of the involved parameters. It is believed that the access to this product representation will greatly enhance the usefulness of CLAAS-3 for some applications.

For the ICDR, which includes data for 2021 and beyond, a backwards-extension (covering the year 2020) was generated exclusively for this Validation Report in order to evaluate the consistency with the TCDR. The evaluation revealed a very high consistency between the ICDR and the TCDR with differences much smaller than the deviations of the TCDR to the reference observations. The ICDR will be subject to annual quality assessments in the future.

The evaluation results included in this report demonstrate the high quality of the products, and indeed show an overall improvement compared to CLAAS edition 2.1 [RD 5]. The quality of the dataset builds on state-of-the-art retrievals systems, on well characterized VIS reflectances and IR radiances, as well as on a carefully implemented processing system and final product generation, which has been established in the European Weather Cloud hosted by EUMETSAT. Further guidance on how to use the products is given in the product user manual [RD 1]. The mentioned accuracy, precision and stability requirements were defined in the product requirements document [AD 1] and were reviewed at the Requirements Review 3.7 [AD 2]. For more details on the used algorithms and aggregation methods the reader is referred to algorithm theoretical basis documents [RD 2 – RD 4].

Table 1-1: Summary of CLAAS-3 validation results compared to requirements for each cloud product. Required and achieved accuracies are formulated in terms of bias, precisions in terms of bc-rmsd (except for CFC and CPH L2, for which the metric is the Hanssen-Kuipers Skill Score, KSS), and stabilities in terms of decadal trend in bias. All numbers, except KSS, are in the units indicated for the respective cloud products. Validation results are color-coded as follows: **worse than threshold**, **fulfils threshold**, **fulfils target**, and **fulfils optimal** requirement. CALIOP results are reported for different values of (I)COT. Evaluations against MODIS are indicated in grey since they are viewed as consistency checks rather than true validation. CPH, LWP and IWP validation scores are for the 3.9- μ m-based products.

L2 or L3	Reference	Accuracy (bias)		Precision (bc-rmsd)		Stability (decadal trend in bias)	
		Req'd (target)	Ach'd	Req'd (target)	Ach'd	Req'd (target)	Ach'd
Cloud Fractional Cover (CFC) [%]							
L2	CALIOP (COT > 0)			0.6 (KSS)	0.67		
	CALIOP (COT > 0.2)				0.70		
L3	SYNOP	5	0.2	10	8.9	2	0.5
	CALIOP (COT>0)		-2.3		22.5		-0.8
	CALIOP (COT \geq 0.3)		5.8		21.0		-1.5
	MODIS		-6.5		9.2		0.3
Cloud Top Height (CTH) [m]							
L2	CALIOP (ICOT > 0)	800	-1015	2400	3032		
	CALIOP (ICOT > 0.2)		230		2260		
	MODIS (liquid)		17		1175		

	Validation Report	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD
	SEVIRI cloud products	Issue: 3.1
	CLAAS Edition 3	Date: 08.08.2022

L2 or L3	Reference	Accuracy (bias)		Precision (bc-rmsd)		Stability (decadal trend in bias)	
		Req'd (target)	Ach'd	Req'd (target)	Ach'd	Req'd (target)	Ach'd
	MODIS (ice)		911		1224		
L3	CALIOP (ICOT > 0)	800	-951	1600	3319	270	178
	CALIOP (ICOT ≥ 0.3)		780		3044		73
	MODIS		2263		1625		-37
Cloud Top Pressure (CTP) [hPa]							
L2	CALIOP (ICOT > 0)	45	42	135	150		
	CALIOP (ICOT > 0.2)		-15		119		
L3	MODIS	45	-167	90	90	15	3.1
Cloud Phase (CPH), defined as fraction of liquid clouds [%]							
L2	CALIOP (ICOT > 0)			0.6 (KSS)	0.70		
	CALIOP (ICOT > 0.2)				0.74		
L3	MODIS	5	-5.9	10	13.2	2	2.4
Liquid Water Path (LWP) [g m ⁻²]							
L2	AMSR2	10	0.1	50	49		
	MODIS		-15		84		
L3	MAC-LWP	10	0.3	20	11.4	3	1.6
	MODIS		-4.0		10.8		0.2
Ice Water Path (IWP) [g m ⁻²]							
L2	DARDAR v3.00	20	-58	100	277		
	DARDAR v3.10		-29		227		
	MODIS		-36		168		
L3	MODIS	20	1.5	40	17.9	6	0.1
Joint Cloud property Histogram (JCH)							
L3		n/a		n/a		n/a	

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

2 The EUMETSAT SAF on Climate Monitoring

The importance of satellite-based climate monitoring 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), the Meteorological Service of the United Kingdom (UK MetOffice) and the Centre National de la recherche scientifique (CNRS) of France. 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 new 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 selected 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, the CM°SAF 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 an international collaboration benefiting from developments at international level and pollinating the partnership with own ideas and standards,

	<p style="text-align: center;">Validation Report SEVIRI cloud products CLAAS Edition 3</p>	<p>Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022</p>
---	---	---

- 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 Programme),
- 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, www.cmsaf.eu. Here, detailed information about product ordering, add-on tools, sample programs and documentation is provided.

3 Introduction to CLAAS-3

This CM°SAF report provides information on the validation of the third edition of the CLOUD property dAtAset using SEVIRI (CLAAS-3), following two earlier editions described in Stengel et al. (2014) and Benas et al. (2017), respectively. All CLAAS-3 products are derived from measurements taken by the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) onboard the EUMETSAT Meteosat Second Generation (MSG) satellites. CLAAS-3 consists of a Thematic Climate Data Record (TCDR), spanning the period from 2004 to 2020, and an Interim Climate Data Record (ICDR), starting in 2021 and extended operationally with low latency into the present (see Figure 3-1). The ICDR is produced with the same algorithms as the TCDR but with a few differences in input data.

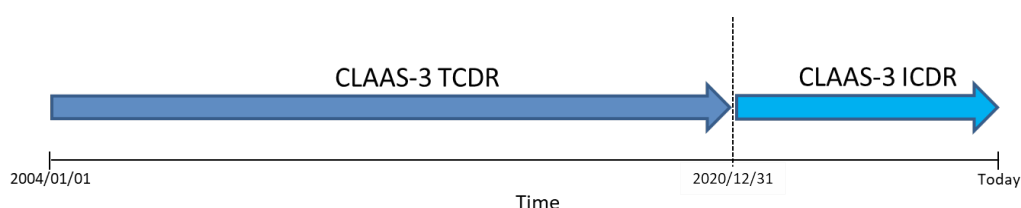


Figure 3-1: Illustration of the CLAAS-3 TCDR and ICDR temporal coverage.

CLAAS-3 consists of six main cloud products, as listed in Table 3-1, with various additional sublayers. These products are provided as instantaneous data (Level 2, L2) and as daily and monthly aggregations (Level 3, L3). A complete overview is given in [RD 1]. The L2 CFC and CTO products were derived using the NWC SAF PPS-v2018-patchCMSAF-May2021 Cloud Probability and Cloud Top Temperature/Height algorithms, which have been adapted to SEVIRI [RD 3] while L2 CPH, LWP and IWP were retrieved with the Cloud Physical Properties (CPP) scheme [RD 4]. The generation of L3 products is described in RD 2]. JCH contains combinations of COT and CTP in histogram form and is composed in the post-processing.

Table 3-1: CLAAS-3 TCDR and ICDR products with formal product IDs and abbreviations according to [AD 1].

Product	TCDR ID	ICDR ID
Fractional Cloud Cover (CFC)	CM-21014	CM-5011
Joint Cloud property Histogram (JCH)	CM-21023	CM-5021
Cloud Top level (CTO)	CM-21033	CM-5031
Cloud Phase (CPH)	CM-21043	CM-5041
Liquid Water Path (LWP)	CM-21053	CM-5051
Ice Water Path (IWP)	CM-21063	CM-5061

The basis of CLAAS-3 is formed by SEVIRI measurements. SEVIRI is a passive visible and infrared imager mounted on the Meteosat Second Generation satellites 1, 2, 3 and 4. MSG1-4, also referred to as Meteosat-8 to -11, are geostationary satellites which, by their rotation,

support a SEVIRI imaging repeat cycle of 15 minutes. SEVIRI itself is an optical imaging radiometer with 12 spectral channels ranging from the visible (approximately 0.6 μm) to the infrared at about 13.4 μm . More details about these channels and the SEVIRI measurements (Level 1.5, L1.5) can be found in [RD 2] as well as in Appendix A.

The respective MSGs in operational mode are centred near 0°/0° latitude/longitude, where a full earth disk image includes Europe, Africa, the Middle East and the Atlantic Ocean. The actual satellite positions are shown in Figure 3-2. L2 products were generated on the native SEVIRI grid, which has a spatial resolution of 3 x 3 km² at nadir and increases towards the edge of the disk (see Figure 3-3). The TCDR covers the time-span 2004-2020, where measurements of the actual operational MSG satellite were processed: MSG1 from 2004-01-19 to 2007-04-11, MSG2 from 2007-04-11 to 2013-01-21, MSG3 from 2013-01-21 to 2018-02-20, and MSG4 from 2018-02-20 to 2020-12-31. Data gaps (e.g., due to sensor decontamination) were filled using data from the backup satellite, if available (see Figure 3-2).

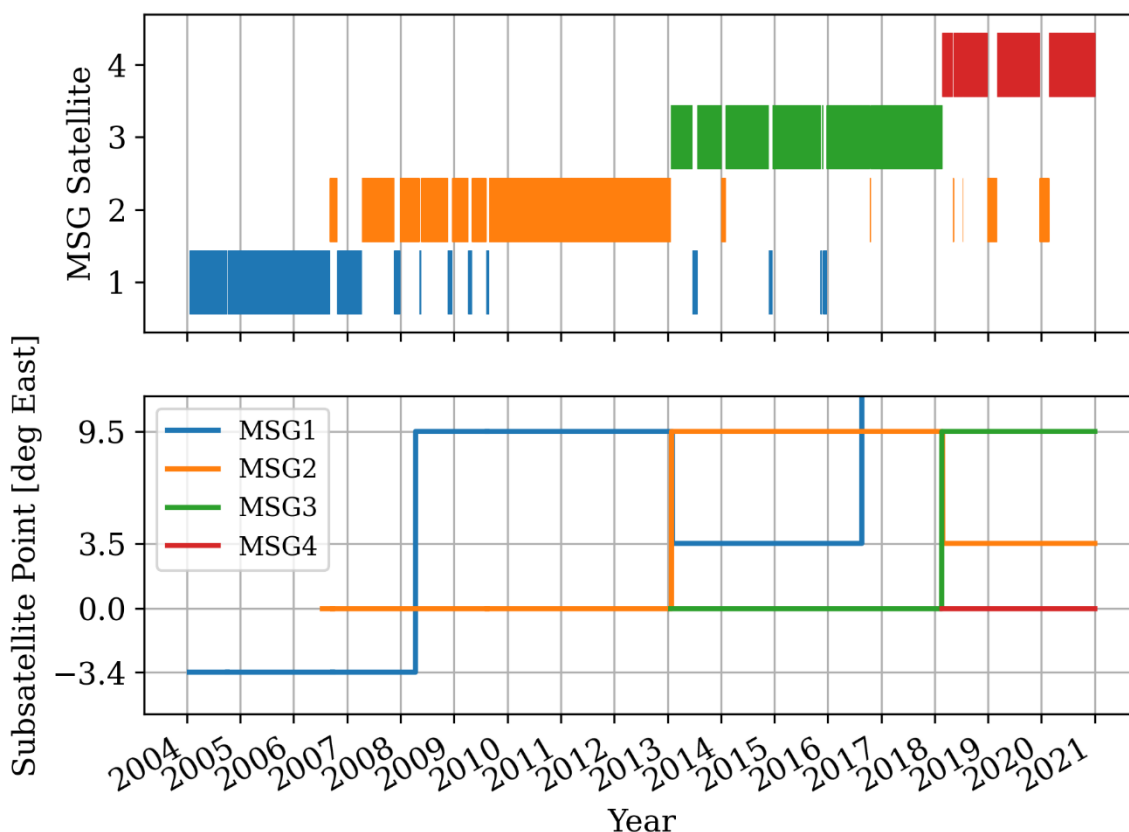


Figure 3-2: Overview of the SEVIRI measurement record used as basis for the generation of the CLAAS-3 TCDR. The top panel shows which satellites were used with short-term data gaps enlarged by a factor 5 for better visibility. The bottom panel shows the location of the satellites along the equator.

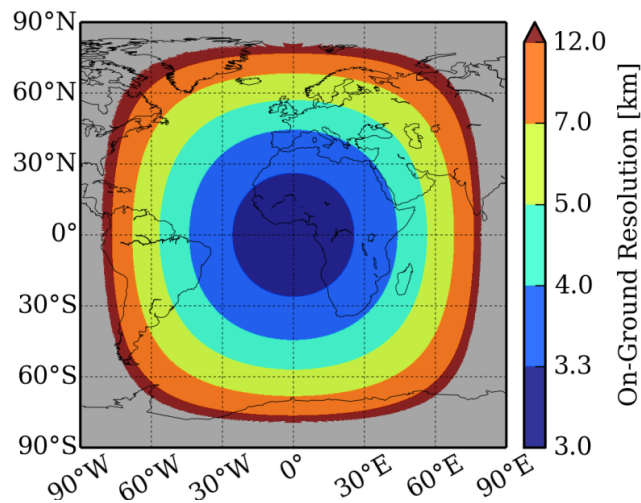


Figure 3-3: SEVIRI spatial coverage and on-ground resolution expressed as the edge length of a square having the same area as the SEVIRI grid cell.

For the ICDR, starting in 2021, the input remains the prime operational MSG satellite near 0°/0° latitude/longitude. Exclusively for this Validation Report, a backwards-extension of the ICDR (covering the year 2020) was generated to evaluate the consistency with the TCDR. The ICDR will be subject to annual quality assessments in the future.

The remainder of this report is organized as follows. First, the overall validation strategy is outlined in Section 4. Then the main changes in CLAAS-3 products compared to the previous edition CLAAS-2.1 are illustrated in Section 5. The validation and evaluation with other datasets are presented in Section 6 for L2 and Section 7 for L3. Verification of the consistency between the ICDR and the TCDR is done in Section 8. Finally, conclusions are drawn in Section 9.

4 Validation strategy

The purpose of the validation effort is to characterize the cloud products in terms of accuracy, precision and stability, thus giving a guidance for applicability of the products. Furthermore, the products are confronted with the product requirements stated in [AD 1] and their compliance is reported.

For geophysical quantities at L2, such as cloud top height, and aggregated (L3) products, we use the bias, i.e. mean difference between CLAAS-3 and reference data as the metric for accuracy. In addition, the bias-corrected root mean squared difference (bc-rmsd) is used to express the precision of CLAAS compared to a reference dataset. In case of discrete L2 variables with only two possible events, e.g. cloud mask (*clear* or *cloudy*) and cloud phase (*liquid* or *ice*), a number of scores are defined which can be derived from a contingency table (Table 4-1).

Table 4-1: Contingency table for the 2x2 problem. n_{ij} is the number of cases where CLAAS-3 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
CLAAS-3 reports 1	n_{11}	n_{12}
CLAAS-3 reports 2	n_{21}	n_{22}

The following scores based on the contingency table are used in this report.

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

These scores can be viewed as measures of precision.

Table 4-2 gives the target requirements for all CLAAS-3 cloud products. Observe that two versions for the Cloud Top Level product (CM-21033) are listed: cloud top height and cloud top pressure. In addition, there are no specific requirements given for the JCH product since it is composed of individual products COT and CTP, which are validated independently. Table 4-2 only lists the target requirements for the accuracy and precision parameters. Compliance with more relaxed threshold requirements and more demanding optimal requirements (as defined in [AD 1]) are also discussed in this report.

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 (2011). 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].

Table 4-2: CM^oSAF CLAAS-3 products and their respective target requirements for accuracy, precision and stability (defined in [AD 1]). L2 requirements refer to the pixel level, L3 requirements to daily/monthly means. The requirements are applicable to the SEVIRI disk.

Product	Accuracy requirement	Precision requirement	Stability requirement
Cloud Fractional Cover (CFC)	L3: bias < 5 % (absolute)	L3: bc-rmsd < 10 % (absolute) L2: KSS > 0.6	stab. of bias < 2%/dec
Cloud Top Height (CTH)	L2/L3: bias < 800 m	L3: bc-rmsd < 1600 m L2: bc-rmsd < 2400 m	stab. of bias < 270 m/dec
Cloud Top Pressure (CTP)	L2/L3: bias < 45 hPa	L3: bc-rmsd < 90 hPa L2: bc-rmsd < 135 hPa	stab. of bias < 15 hPa/dec
Cloud Phase (CPH)	L3: bias < 5% (liquid cloud fraction, absolute)	L3: bc-rmsd < 10 % (liquid cloud fraction, absolute) L2: KSS > 0.6	stab. of bias < 2%/dec
Liquid Water Path (LWP)	L2/L3: bias < 10 g m ⁻²	L3: bc-rmsd < 20 g m ⁻² L2: bc-rmsd < 50 g m ⁻²	stab. of bias < 3 g m ⁻² /dec
Ice Water Path (IWP)	L2/L3: bias < 20 g m ⁻²	L3: bc-rmsd < 40 g m ⁻² L2: bc-rmsd < 100 g m ⁻²	stab. of bias < 6 g m ⁻² /dec
Joint Cloud Histogram (JCH)	n/a	n/a	n/a

The CLAAS-3 TCDR consists of instantaneous (L2) data and daily/monthly mean (L3) products for the period 2004-2020. Furthermore, monthly mean diurnal cycles are composed as well as monthly single- and multi-parameter histograms. The validation task comprises evaluation of the L2 and L3 products, with the latter done if reference measurements were available. However, inter-comparisons with L3 products from other sources are often more difficult to interpret than comparisons with instantaneous and simultaneous observations (i.e., the classical L2 validation process). The reason is that L3 products not only depend on the quality of L2 products but also on the method of compiling L3 products (i.e., in terms of the applied temporal and spatial sampling, criteria for including or excluding a measurement, averaging

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

method, etc.). This means that L3 product differences do not always reflect true product differences in the same way as monitored by standard L2 validation activities.

For practical reasons L2 studies have been limited in time compared to the task of evaluating the full CLAAS-3 dataset. In order to keep data volumes and processing time within reasonable limits, it was decided to concentrate part of the L2 validation efforts on a single month, namely March 2013, for which all required reference instruments and datasets are available. This allows adequate and in-depth characterization of the L2 products. Obviously, variations on seasonal and multi-annual timescales cannot be assessed in this way. However, these are covered by the L3 evaluation presented in Section 7. We believe that the mix of L2 (instantaneous) and L3 (monthly mean) studies provides sufficient information about the expected quality of daily L3 products, which were not separately evaluated. The evaluation in this report is done with respect to 'best practice', using well-established, high quality, homogeneous and independent observations, to the extent that these are available.

The chosen validation references may be subdivided into two groups:

- **independent observations**, which are generally considered to be true references, i.e. of superior quality. We have used the following observations:
 - Cloud amount from surface stations (SYNOP), 2004-2020 (L3)
 - Cloud amount, cloud phase, and cloud top level from space-based lidar CALIOP: 2013 (L2) and 2006-2016 (L3)
 - Ice cloud properties from space-based lidar+radar DARDAR: March 2013 (L2)
 - Liquid water path from combined passive microwave sensors: 2004-2016 (L3)
 - Liquid water path from passive microwave sensor AMSR2: March 2013 (L2)
- **similar satellite datasets** based on passive VIS-IR measurements, which are used for inter-comparisons rather than pure validation
 - All cloud properties from MODIS: March 2013 (L2) and 2004-2020 (L3)

The first group of observations is 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, this type of reference observation is not available for all considered parameters. It also has to be realized that none of these independent observation types is perfect, and they all have their uncertainties related to, for example, representativeness, sensitivity, and aggregation. The reference datasets, including associated uncertainties, are described in detail in Appendix B.

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

5 Brief characterization of systematic changes from CLAAS-2.1 to CLAAS-3

This section shows the differences between the cloud properties from the previous CLAAS-2.1 and the new CLAAS-3 data record. First, L3 monthly mean products are compared using the time series and maps of temporally averaged cloud parameters to demonstrate their homogeneity and large-scale patterns. Then, differences in diurnal cycles are shown to illustrate day-night transitions as well as cloud bow and glory impacts. Finally, histograms of L2 cloud properties and occurrence frequencies of failed CRE retrievals over the full SEVIRI disk are compared.

5.1 Time series of L3 products

For the time series the latitude-weighted spatial averages of monthly mean variables were computed. The averages were calculated over all pixels excluding the values north and south of 60° latitude to avoid the influence of polar night conditions. For the comparisons a subset of variables from each L3 monthly mean product is taken.

CLAAS-3 CFC is up to 6% lower than CLAAS-2.1 CFC over the whole observation period. Figure 5-1 shows that cloud fractional cover in CLAAS-2.1 is more stable over time. The difference plot in the right panel reveals a modest increase between 2010 and 2013 in the globally averaged CLAAS-3 CFC. Looking at the daytime and nighttime CFC separately, one can see that the increase is only present at night, which suggests a relation with the IR channels. Indeed, the anomalies may be linked to potential calibration issues in the IR channels, as reported in Appendix A. Specifically, an irregularity can be seen in the 3.9 – 10.8 μm BT difference in about the same years (Figure 10-2), although anomalies in other channels may contribute as well. The daytime CFC difference between the CLAAS-3 and CLAAS-2.1 remains stable at a value of approximately -11% over the whole overlapping period, while the nighttime difference is, apart from the irregularities, only -2%.

The global average of CPH, the fraction of liquid water clouds, is more stable in CLAAS-3 than in the previous edition (Figure 5-2). A jump in CPH in January 2013, caused by the switch from Meteosat-9 to Meteosat-10, and a slightly positive trend from 2013 onwards have been removed in CLAAS-3. Overall, CPH and daytime CPH in CLAAS-3 show somewhat lower values than in CLAAS-2.1, especially in the last decade.

The time series of cloud top parameters, height, pressure and temperature, in Figure 5-3, show similar behaviour and a homogeneous difference between both editions of CLAAS before 2013. Similar to the CPH time series, the jump in 2013 has been removed in CLAAS-3, and the difference between the data records becomes larger afterwards. On average, CLAAS-3 contains higher and colder clouds.

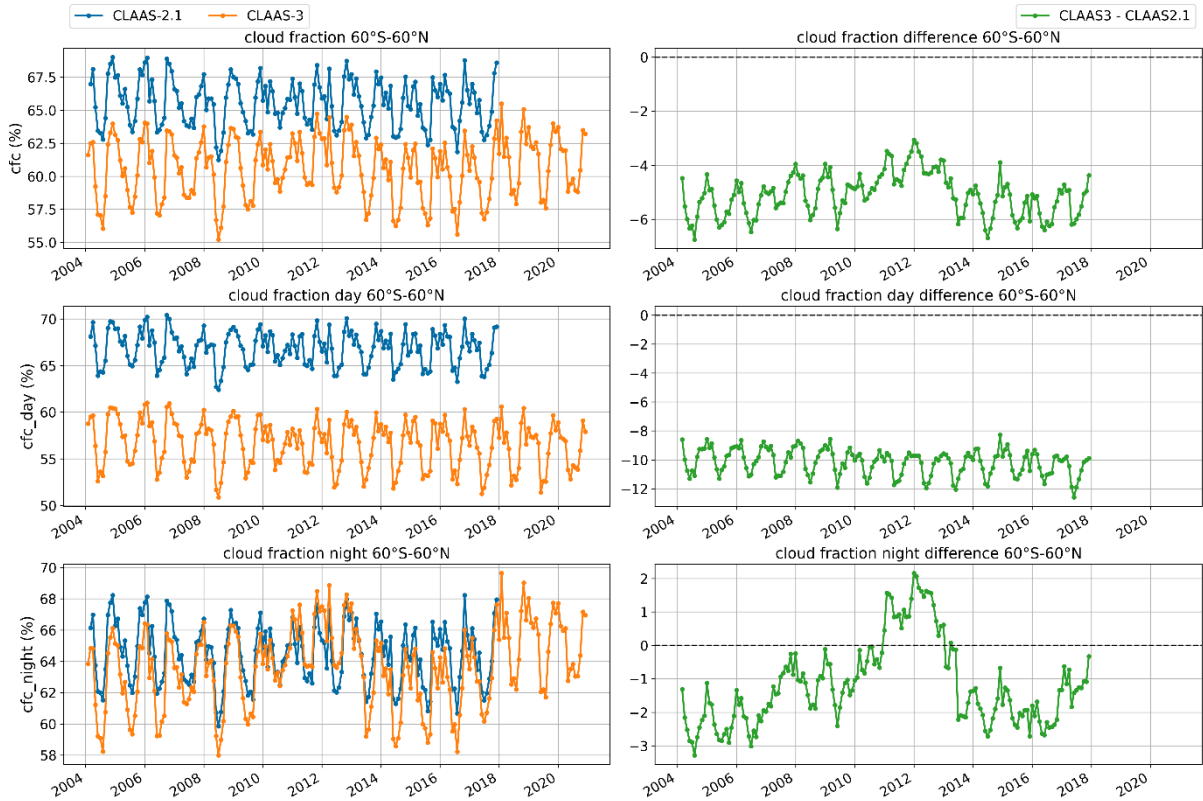


Figure 5-1: Time series of the CLAAS-2.1 and CLAAS-3 CFC products (left column) and time series of the difference between both data sets (right column).

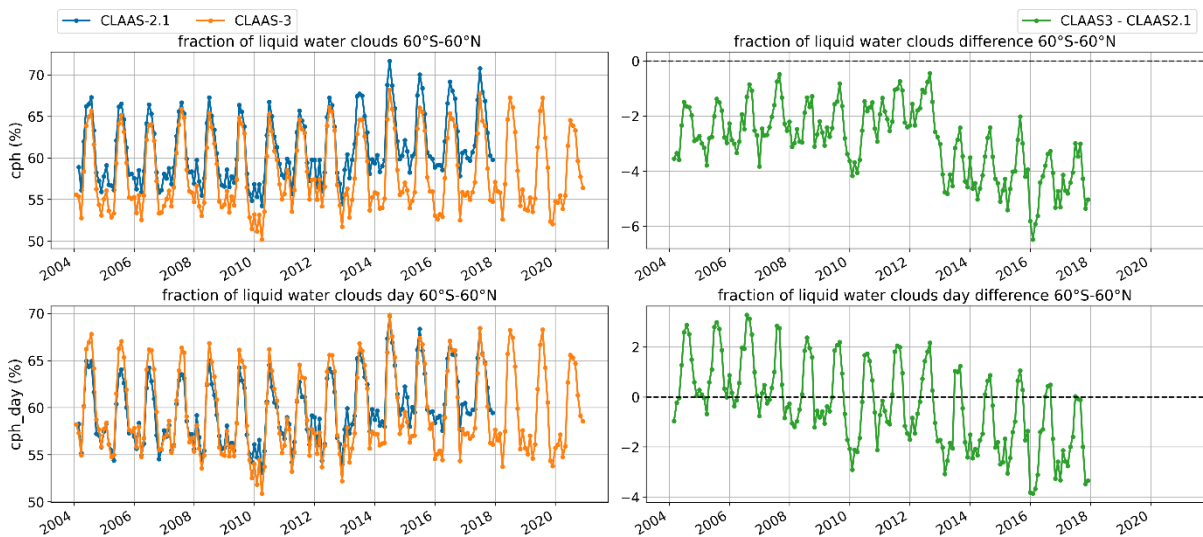


Figure 5-2: Time series of the CLAAS-2.1 and CLAAS-3 CPH products (left column) and time series of the difference between both data sets (right column).

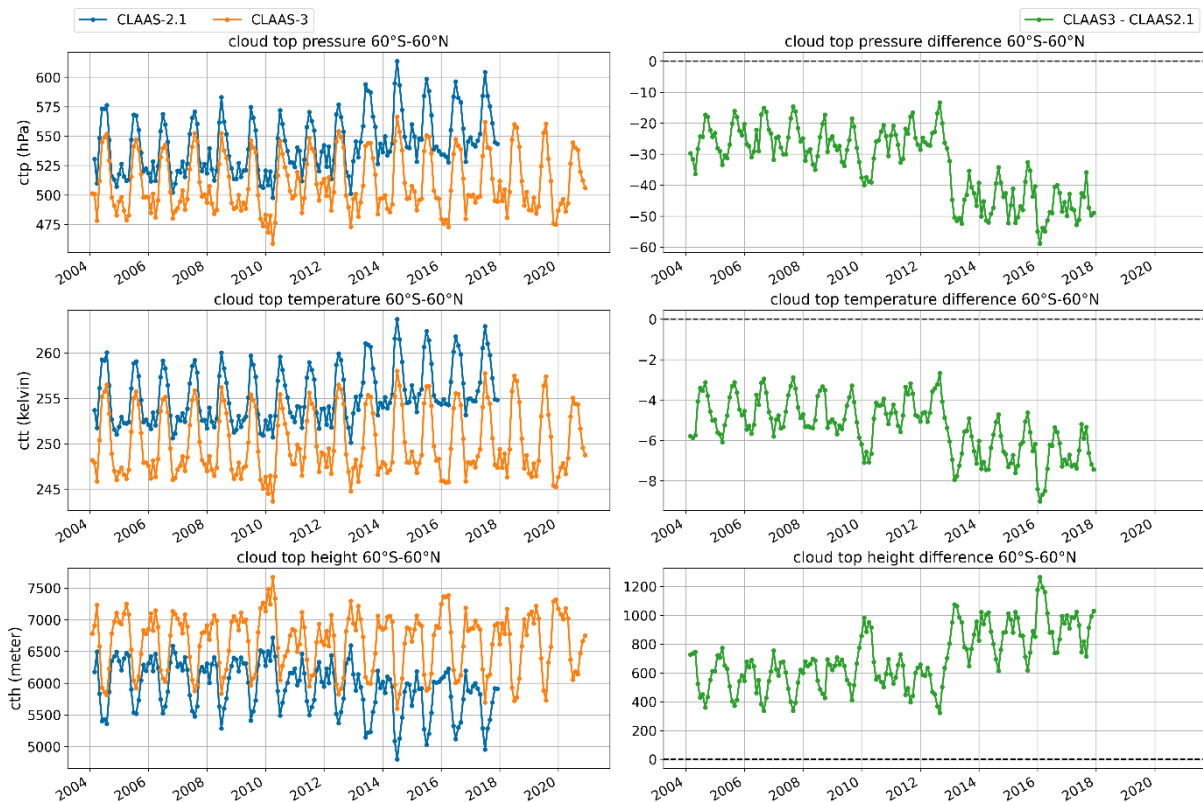


Figure 5-3: Time series of the CLAAS-2.1 and CLAAS-3 CTO products (left column) and time series of the difference between both data sets (right column).

CLAAS-3 contains two sets of LWP and IWP products, using either the 1.6- μm or the 3.9- μm channel for the COT-CRE retrieval. Since CLAAS-2.1 only used the 1.6- μm channel, the comparisons of LWP in Figure 5-4 and IWP in Figure 5-5 are for the CLAAS-3 products based on that same channel. CLAAS-3 shows about 15% higher COT than CLAAS-2.1 for both liquid and ice clouds. This is partly related to the lower daytime CFC, with predominantly (very) thin or false clouds having been removed compared to CLAAS-2.1, as well as to the higher maximum COT, which was increased from 100 to 150. For CRE the comparison is more complicated since in CLAAS-2.1 the CRE assigned to observations outside the forward model solution space was included in the L3 products, whereas this is not the case in CLAAS-3. As shown in Section 5.3, this concerns a significant fraction of the pixels, so it can influence the monthly mean values considerably. While this is the main factor explaining the differences in liquid CRE, for ice clouds it is mainly the adoption of a new ice particle model causing increases in CRE (see also Section 5.3). The all-sky LWP and IWP are driven by the respective COT and CRE but also by CPH.

A change in all-sky LWP around 2007 can be related to the shift of the satellite position after the change from MSG1 to MSG2 (Figure 3-2). An increase of all-sky LWP in the last three years of the CLAAS-2.1 data record is not noticeable in CLAAS-3, so the time series of all-sky LWP in the new edition is more homogeneous in the last years. All-sky IWP as well as its difference between both CLAAS editions is stable over the observation period. A slight negative trend in liquid COT in CLAAS-2.1 is not recognisable in CLAAS-3 as one can see in the increasing difference in the right panel of Figure 5-4. CLAAS-3 ice COT shows a slight

decrease and less seasonal variation after 2007; in CLAAS-2.1 ice COT is stable over the whole observation period. Effective radius of liquid and ice cloud droplets in CLAAS-3 and CLAAS-2.1 are remarkably stable.

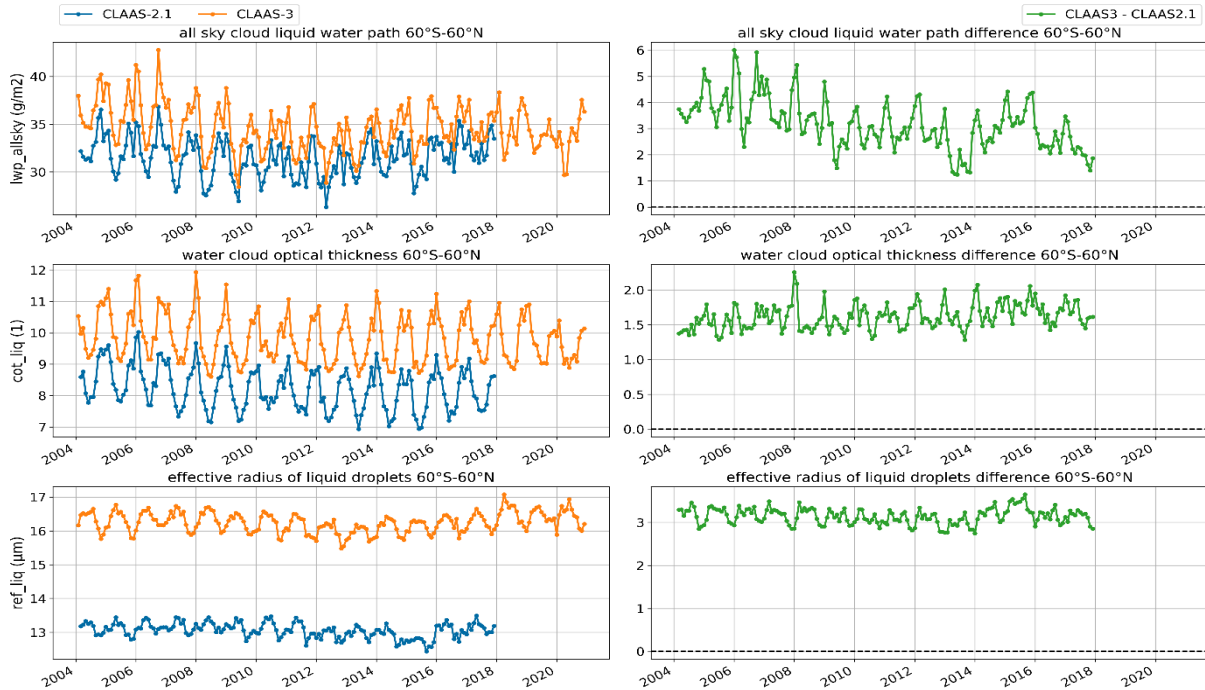


Figure 5-4: Time series of the CLAAS-2.1 and CLAAS-3 LWP products (left column) and time series of the difference between both data sets (right column).

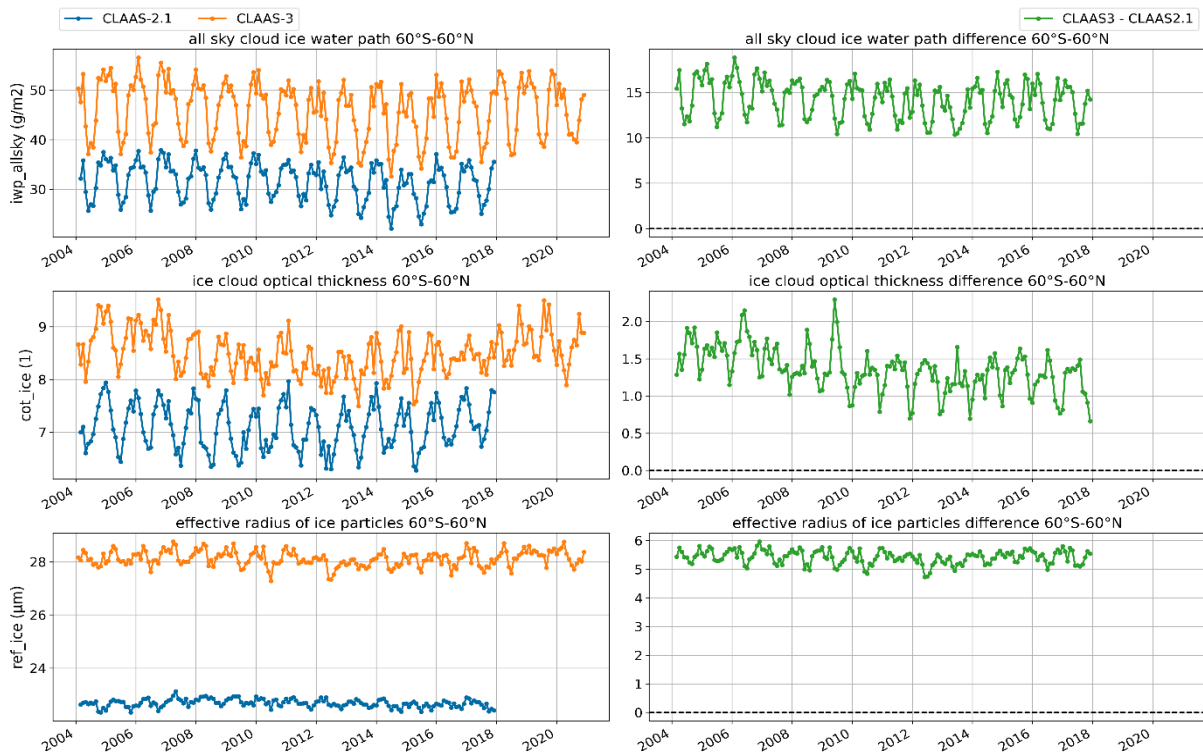


Figure 5-5: Time series of the CLAAS-2.1 and CLAAS-3 IWP products (left column) and time series of the difference between both data sets (right column).

5.2 Spatial maps of L3 products

This section presents a comparison of spatial differences of L3 monthly mean products. The maps in Figure 5-6 have been calculated by temporal averaging of the monthly means of CFC, CTP, CPH, all-sky LWP and all-sky IWP.

All analysed cloud properties from CLAAS-2.1 and CLAAS-3 show similar large-scale patterns. For all variables in CLAAS-3 remarkably higher values at the edge of the disk are noticed.

CLAAS-3 CFC shows lower values over the ocean and for high satellite viewing angles. As mentioned before, at the edge of the disk CLAAS-3 has very high cloud fractions. The difference in CFC between the CLAAS editions is lowest over the African continent and the Arabian Peninsula, where the average cloud fraction is relatively low. The difference is also low at high latitudes with very high average cloudiness.

The difference patterns between CLAAS-3 and CLAAS-2.1 in CTP and CPH are similar: CLAAS-3 contains overall lower values (i.e. higher clouds and larger fraction of ice clouds) except in areas with very high viewing angles. There are no significant changes in the marine stratocumulus region in the Southeast Atlantic. CLAAS-3 shows higher CTP and CPH values over the northernmost part of Africa and the Arabian Peninsula (but note that the cloud fraction is very low there), whereas the African tropics and the ITCZ have more high ice clouds, with up to 20% higher ice cloud fraction and up to 200 hPa lower CTP on average.

	<p style="text-align: center;">Validation Report SEVIRI cloud products CLAAS Edition 3</p>	<p>Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022</p>
---	---	---

All-sky liquid water path in CLAAS-3 and CLAAS-2.1 shows very small changes in the tropics. Near the poles and over land in the mid-latitudes CLAAS-3 all-sky LWP is higher than in CLAAS-2.1. The only area where the averaged all-sky LWP in the new data set edition is lower is the African coast near the equator. All-sky IWP shows greater differences between the dataset editions. The ice water path near the equator is over 40 g/m² higher in CLAAS-3. In the mid- and high-latitudes CLAAS-3 has up to 30 g/m² higher values.

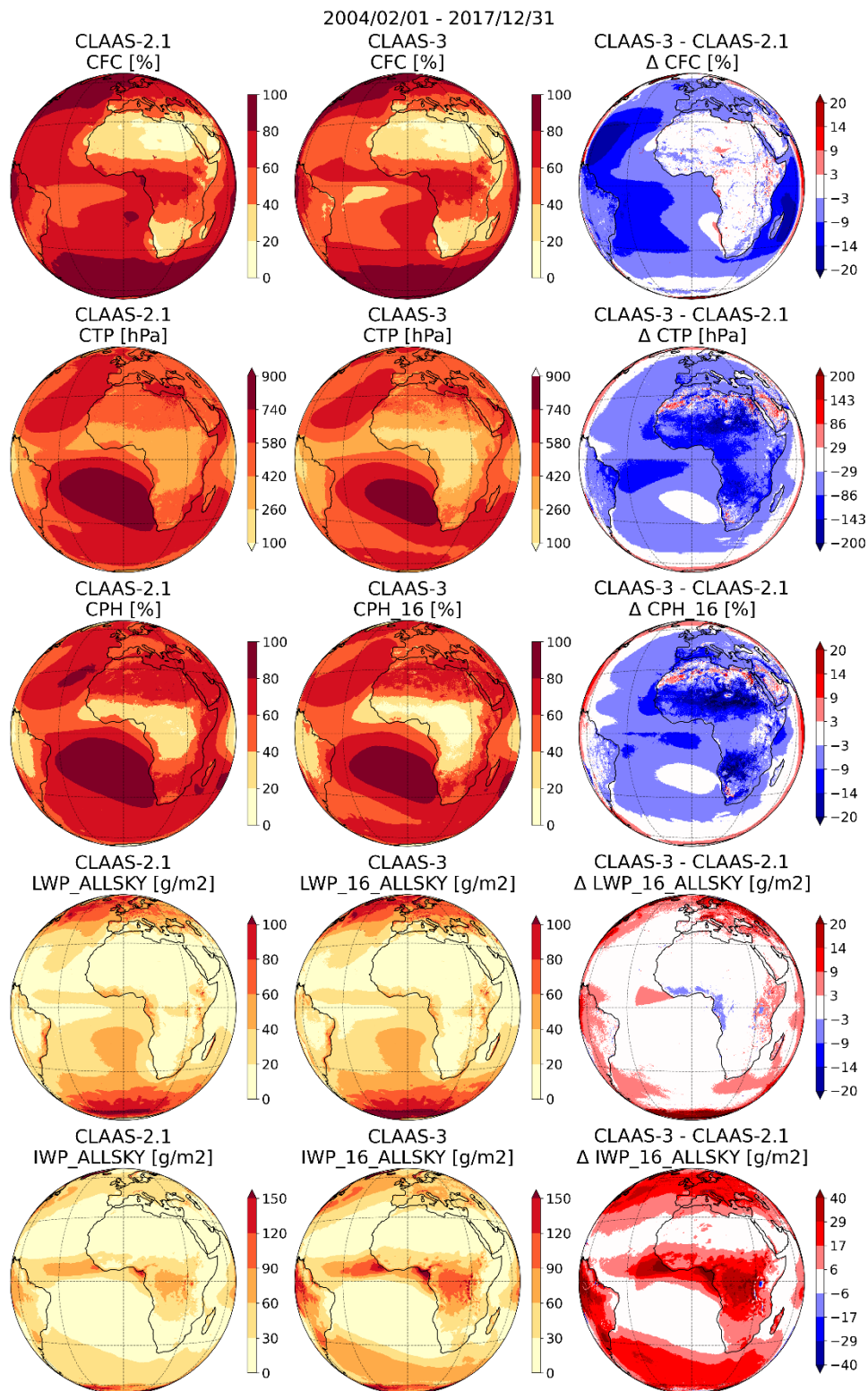


Figure 5-6: Maps of averaged cfc, cph, ctp, lwp_allsky, and iwp_allsky from CLAAS-2.1 (left column), CLAAS-3 (middle column) and the corresponding differences (right column) for the overlapping period: 01/02/2004 – 31/12/2017.

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

5.3 Some characteristics of L2 products and their diurnal cycle

Significant changes have been made between the cloud property retrieval algorithms in CLAAS-3 compared to CLAAS-2.1. Here we highlight some of the consequences for the cloud products, in particular aspects not covered by the larger-scale L3 comparisons in the previous subsections.

The diurnal cycle of cloud property retrievals can show features related to the changing position of the sun rather than to actual physical variations. Such features are typically hidden or averaged out in L3 data. Therefore, diurnal cycles over a relatively small region in the southeast Atlantic, similar to the region used in Benas et al. (2019), were analyzed. Consistently observed patterns are illustrated on the basis of results for a typical day, 25 March 2013, with near overcast conditions during day- and nighttime in Figure 5-7. While this gives an indication of typical changes between the CLAAS-3 and CLAAS-2.1 products and can expose irregularities and artefacts, there are no reference data available to judge their respective quality. Such judgements will be made in the following sections.

CLAAS-3 CFC is similar to CLAAS-2.1 during nighttime and up to 20% lower during daytime. A CFC decrease by 5-15% during twilight conditions is seen in both CLAAS-2.1 and CLAAS-3. CLAAS-3 has higher cloud tops than CLAAS-2.1. Reference data are not available for this particular situation, but validation results (e.g., Figure 6-10) show that CLAAS-3 corresponds generally well to CALIPSO observations for cases with low-level liquid clouds. CLAAS-3 CTH increases by (in this case, and averaged over the region) about 1000 m at dawn and 500 m at dusk (solar zenith angle around 80 degrees). Further analysis (not shown) reveals that this increase is due to too high CTH values (between 5 and 15 km) being retrieved for about 10% of the pixels in the region. The deviations in CFC and CTH during twilight conditions are most probably artefacts in the retrieval, related to the rapidly changing appearance of the 3.9 μm channel, in combination with the limited amount of data available for training the algorithms under these conditions.

The liquid cloud fraction is 100% in both CLAAS-2.1 and CLAAS-3 throughout day and night, except for a decrease by up to 4% corresponding with the increases in CTH. This decrease is explained by the portion of pixels with very high CTH, and thus low CTT, which are labeled as ice phase. COT is comparable between CLAAS-2.1 and the two flavors of CLAAS-3. It shows a strong decrease in the morning, followed by a stabilization and slight increase in the late afternoon. Small irregularities are visible around the times with cloud bow and glory scattering conditions. CRE is comparable between the CLAAS-2.1 and CLAAS-3 1.6 μm retrievals, and considerably lower in the 3.9 μm retrieval. The 1.6 μm retrievals show strong irregularities during cloud bow and glory conditions. The 3.9 μm retrievals only show a (somewhat smaller and more spread out) irregularity during the glory conditions. The 1.6 μm retrievals are below the LUT during cloud bow conditions and mainly above the LUT around the glory timeslots. The 3.9 μm retrievals are frequently above the LUT during daytime in general and near twilight in particular.

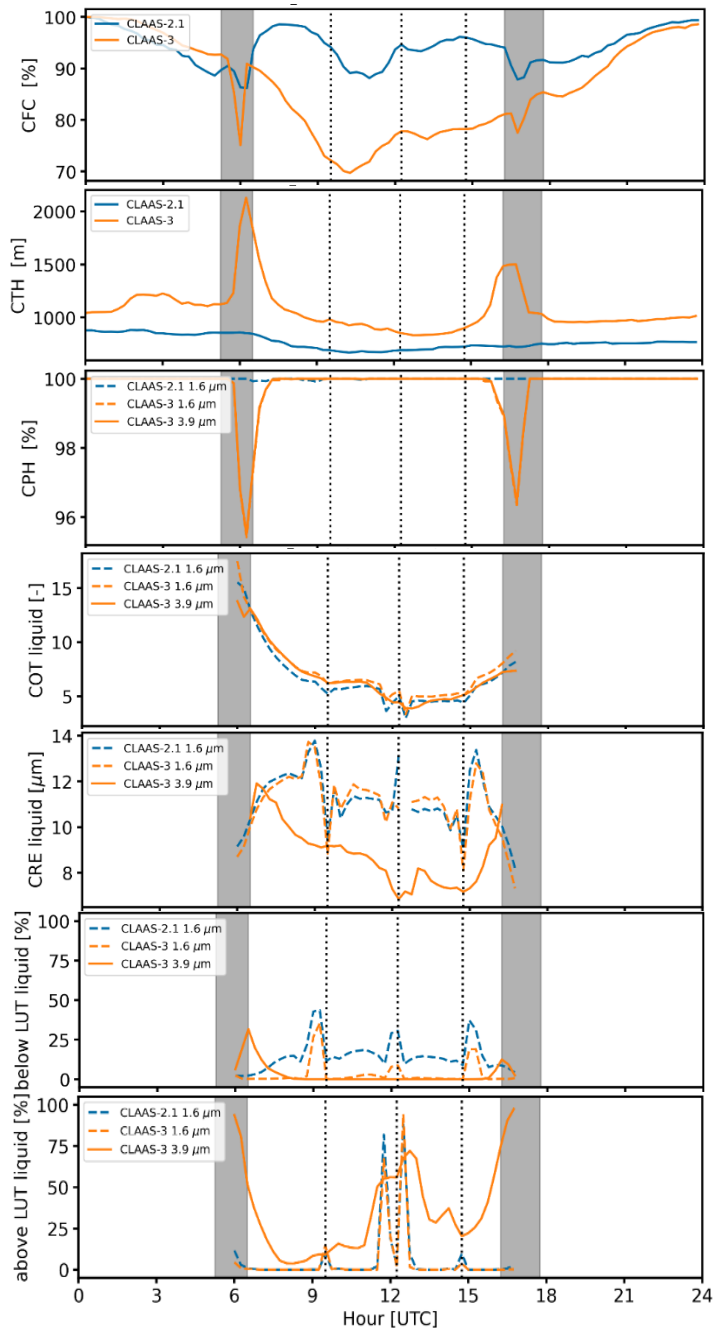


Figure 5-7: Diurnal cycle of cloud CLAAS-2.1 and CLAAS-3 cloud properties averaged over a 100x100 pixel region in the southeast Atlantic on 25 March 2013. Twilight conditions (SZA between 75 and 95 degrees) are indicated by the shaded area. The middle vertical dotted line represents cloud glory conditions (scattering angle near 180 degrees), while the other two vertical dotted lines represent cloud bow conditions (scattering angle near 140 degrees).

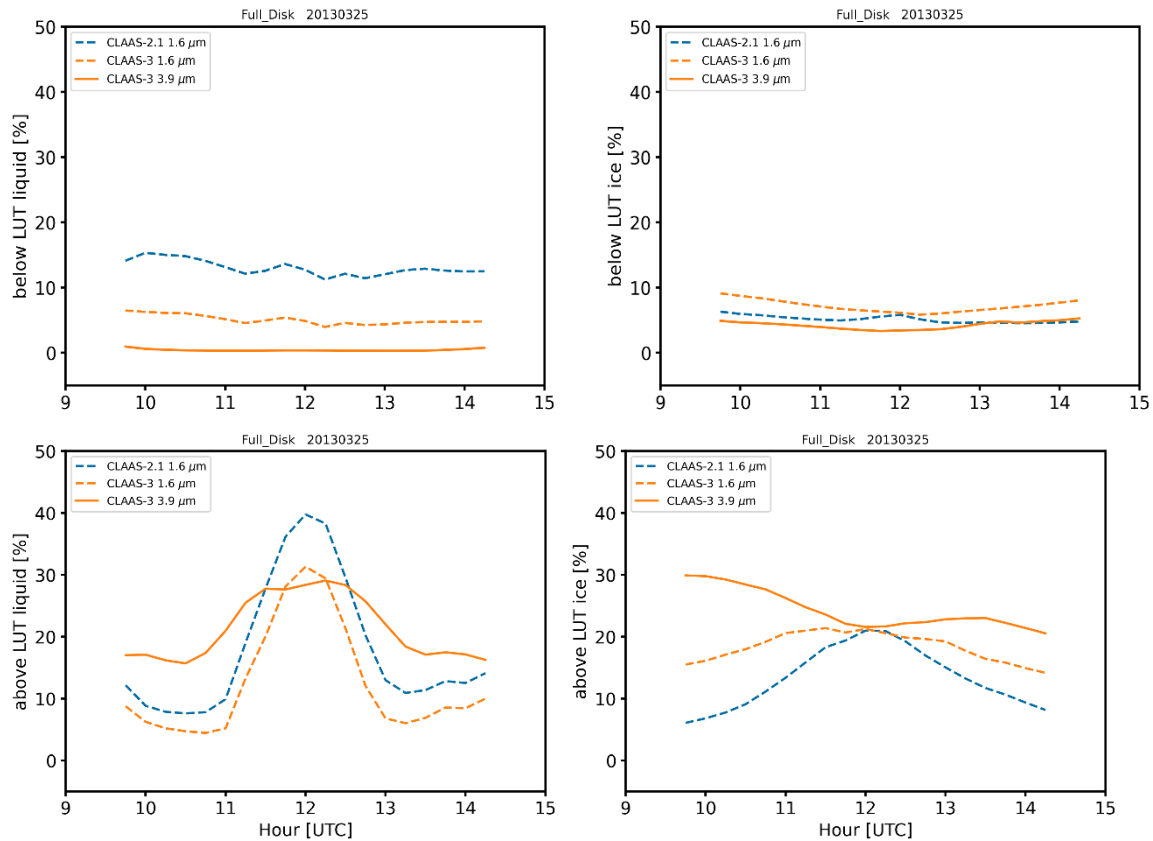


Figure 5-8: Diurnal cycle of the fraction of COT-CRE retrievals below the look-up table (top) and above the look-up table (bottom), for liquid clouds (left) and ice clouds (right) on 25 March 2013. The full disk is considered with the requirement that 70% of the pixels satisfies the retrieval criteria for solar and satellite zenith angles.

A more general indication of the frequency of CRE retrieval failure is given in Figure 5-8, where the full disk is considered. Daily averages are also reported in Table 5-1. One specific day is shown but results do not vary much. CLAAS-3 1.6 μm liquid cloud CRE retrievals are more often inside the LUT than CLAAS-2.1. Between 11 and 13 UTC a large part of the disk is viewed in glory conditions, which explains the increased frequency of above-LUT cases in both CLAAS-3 and CLAAS-2.1. The CLAAS-3 3.9 μm liquid cloud CRE retrievals are hardly ever below the LUT but frequently above. This asymmetry suggests a general imbalance between observed and modelled brightness temperatures in the 3.9 μm channel, which will be a subject of further study. The 3.9 μm liquid cloud retrievals also show an increasing number of above-LUT cases between 11 and 13 UTC, but this feature is weaker than for the 1.6 μm retrievals. For ice clouds the occurrence of below LUT retrievals is quite similar between the products. Above-LUT retrievals occur more often, with the largest occurrence frequency for the CLAAS-3 3.9 μm retrievals. Overall, about 25% of the COT-CRE retrievals fails because no solution can be found inside the look-up table.

To give an impression of changes in L2 cloud properties between CLAAS-2.1 and CLAAS-3, histograms of full disk retrievals are shown in Figure 5-9. Similar histograms are also included in Section 6, but there the purpose is validation and inter-comparison with other datasets, and

the observations are subject to spatio-temporal collocation, spatial aggregation and (in some cases) filtering, which can change the characteristics.

Table 5-1: Percentage of liquid and ice CRE retrievals below and above the look-up table, averaged over the full disk on 25 March 2013.

	Liquid			Ice			Liquid + Ice		
	below	above	total	below	above	total	below	above	total
CLAAS-2.1 1.6 μm	13.0	17.9	30.9	5.1	13.4	18.5	9.6	16.0	25.6
CLAAS-3 1.6 μm	5.1	12.5	17.6	7.1	18.4	25.5	6.0	15.2	21.2
CLAAS-3 3.9 μm	0.4	21.3	21.7	4.2	24.3	28.5	2.1	22.6	24.7

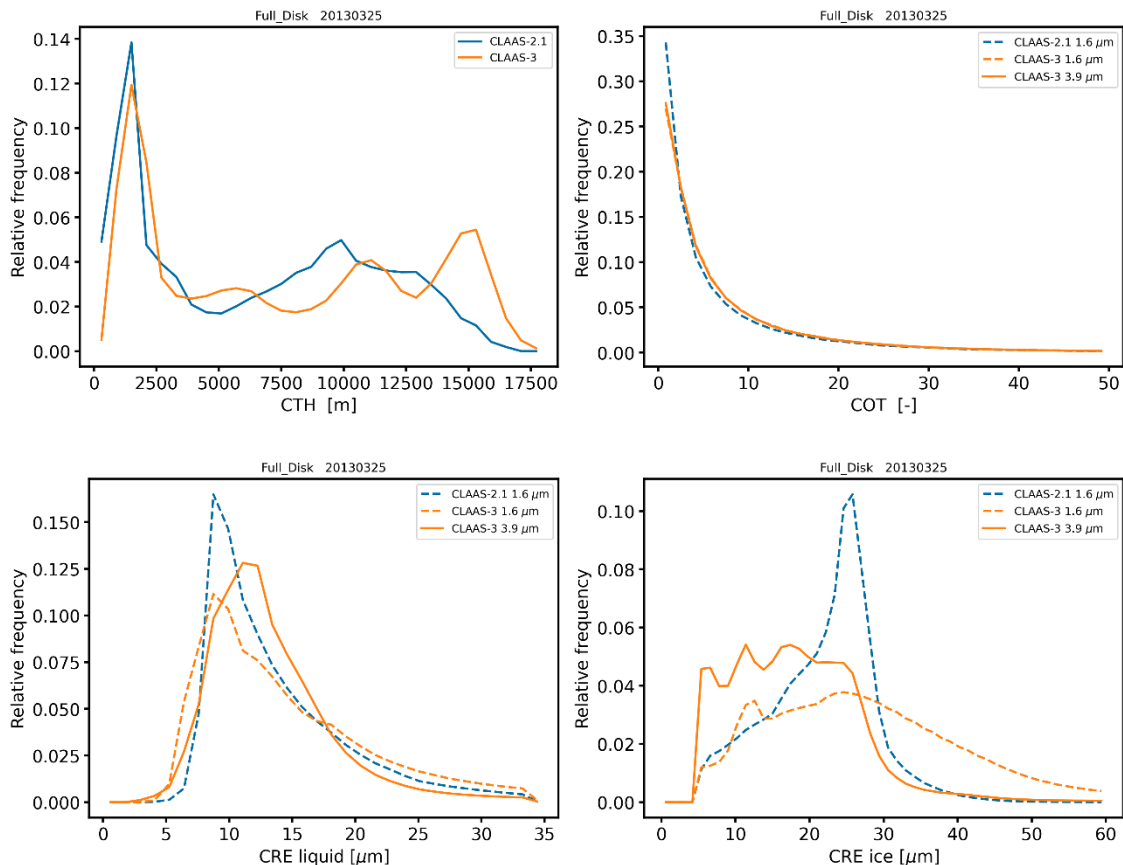


Figure 5-9: Histograms of cloud top height, cloud optical thickness, liquid cloud droplet effective radius, and ice cloud particle effective radius from CLAAS-2.1 and CLAAS-3 full-disk retrievals on 25 March 2013.

Cloud-top height distributions contain several key differences. CLAAS-3 cloud tops below 600 m are nearly absent, whereas they frequently occur in CLAAS-2.1. Cloud tops between about 7 and 10 km height occur less frequently in CLAAS-3, whereas the opposite is true for cloud tops higher than 13 km. Cloud optical thickness histograms are quite similar, with relatively

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

somewhat less thin clouds retrieved in the two CLAAS-3 flavors compared with CLAAS-2.1. Liquid cloud effective radii in the 1.6 μm channel have a wider distribution in CLAAS-3 than in CLAAS-2.1. The main reason for this is that the weighting with a climatological/prior effective radius (set at 8 μm) for thin clouds in CLAAS-2.1 was dropped in CLAAS-3. Other changes, notably a decrease in the assumed droplet size distribution width, play a role as well. The 3.9 μm CRE retrievals are characterized by a somewhat larger mode and shorter tail than the 1.6 μm retrievals. For ice clouds the various CRE retrievals show larger differences. CRE from the 1.6 μm channel has a much wider distribution in CLAAS-3 than in CLAAS-2.1, which is again related to removal of the weighting with a climatological value (of 26 μm) for thin clouds but even more to the change in assumed ice particle habits. The 3.9 μm ice CRE has rather low values, frequently near the minimum retrievable value of 5 μm and rarely exceeding 30 μm .

5.4 Summary of differences between CLAAS-2.1 and CLAAS-3

In summary, the main differences between CLAAS-2.1 and CLAAS-3 presented in this section are the following:

- CLAAS-3 has a lower CFC, in particular during daytime and over ocean.
- CLAAS-3 CTO is retrieved for nearly all cloudy pixels, which was not the case for CLAAS-2.1, and it shows overall higher cloud-tops
- CLAAS-3 has relatively more ice clouds as well as a larger IWP.
- CLAAS-3 nighttime CFC time series show positive anomalies of a few percent in the period 2010-2013.
- A discontinuity in CLAAS-2.1 CPH and CTO time series in January 2013 is not present anymore in CLAAS-3.
- CLAAS-3 has too high CFC in a narrow zone near the edge of the SEVIRI disk.
- CLAAS-3 CTO products show positive anomalies for low clouds in the diurnal cycle around twilight.
- CLAAS-3 COT and CRE retrievals show irregularities for specific viewing and illumination geometries including cloud bow and glory conditions, which were also present in CLAAS-2.1.
- CLAAS-3 (1.6- μm based) CRE shows wider distributions at L2. CLAAS-3 CRE at L3 is larger for both liquid and ice clouds, resulting from various changes in the retrieval and aggregation process.

6 Evaluation of CLAAS-3 instantaneous (L2) cloud parameters

This section covers the evaluation of CLAAS-3 L2 products and is organized according to Table 6-1.

Table 6-1: Overview of reference datasets used for the evaluation of CLAAS-3 L2 parameters.

Section	Reference observations	Parameters
6.1	Calipso	CFC, CPH, CTH
6.2	DARDAR (CloudSat-Calipso)	IWP, ice COT, ice CRE
6.3	AMSR2	LWP
6.4	MODIS	CTH, COT, CRE, CDNC, CGT, LWP, IWP

6.1 Validation with CALIOP

The validation against CALIOP on CALIPSO requires the spatial and temporal matching of observations from SEVIRI and CALIOP. Collocations were computed using spatial nearest neighbour search and scanline-based time matching: For each CALIOP measurement, the nearest neighbour in the SEVIRI grid is determined yielding its line number in the image. The matching SEVIRI scan is given by the scan where the acquisition time of this particular scanline is closest to the timestamp of the CALIOP measurement. An example matchup is shown in Figure 6-1. Maximum collocation distances are 5 km and 7.5 minutes in space and time.

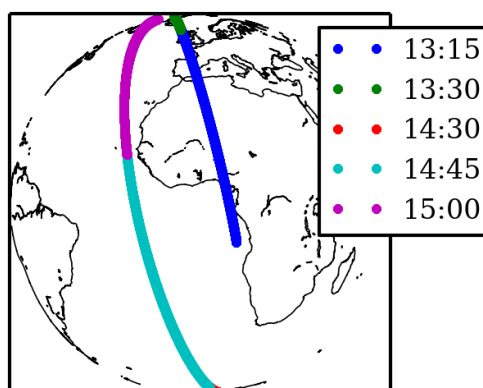


Figure 6-1: Example matchup of CALIOP track and SEVIRI scan. The labels display the nominal timestamp of the SEVIRI scan that was matched to the identically coloured ground track segment.

Due to the advanced lidar technique, CALIOP is much more sensitive to high and optically thin clouds than SEVIRI. Therefore, we do not only compare CLAAS-3 against the uppermost cloud layer detected by CALIOP, but also compare against CALIOP data which was filtered by means of the cloud optical thickness (COT). The latter can tell us more about how accurate CLAAS-3 products are relative to the potential of the SEVIRI sensor.

Collocations were made for the complete year 2013, limiting the amount of data but ensuring that all seasons are covered. The year 2013 was not part of the dataset used to train the probabilistic cloud mask and cloud top temperature/height algorithms [RD 3].

In the following validation sections, we excluded measurements known to be of reduced quality:

- Absolute value of latitude/longitude > 80 degrees.
- SEVIRI satellite viewing zenith angle (VZA) > 75 degrees.
- CALIOP phase information flagged as *bad quality*.

6.1.1 Cloud Mask and Cloud Phase

After the collocation of CLAAS-3 and CALIOP the cloud fraction of CALIOP was binarized using a 50% cloudiness threshold. This is required to allow for a comparison against the CLAAS-3 binary cloud mask. The cloud phase from CALIOP distinguish different ice crystal orientations which were combined to *ice phase* for direct comparison to CLAAS-3 including only liquid and ice phases.

CALIOP's cloudy measurements correspond to the total column cloud optical thickness (COT) > 0. For additional sensitivity analyses for CLAAS-3 cloud detection we also include comparisons against CALIOP cloud masks which were derived using total column COT threshold larger than zero. A reasonable approximate COT-threshold to use, matching the detection limit for clouds interpreted from passive imagery, is 0.2 according to the study by Karlsson and Håkansson (2018) and . It is important to note, that in this scenario CALIOP data with COTs between 0.0 and 0.2 are not excluded from the statistics but rather re-classified as clear-sky.

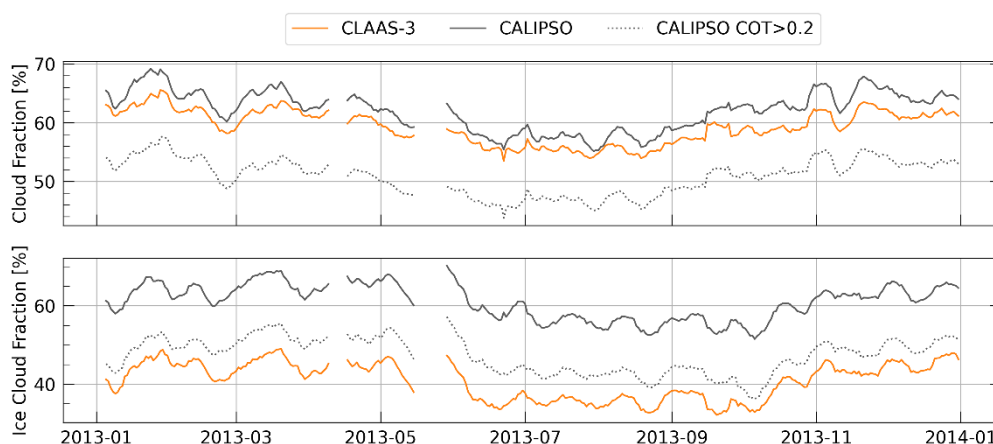


Figure 6-2: 10-day moving average of cloud mask (yielding a cloud fraction, top) and cloud phase (bottom, yielding an ice cloud fraction) from CLAAS-3 and CALIOP, averaged over the full disk. Dashed lines show COT-filtered CALIOP data.

For cloud phase the additional sensitivity analysis does not consider the total column COT, but instead selects the cloud phase reference from CALIOP profiles by excluding the uppermost cloud layers until the integrated layer-COT (top down) exceeds the selected threshold of 0.2 (ICOT = 0.2 hereafter).

Figure 6-2 shows the time series of cloud fraction and ice cloud fraction from CLAAS-3 and CALIOP at all collected collocations in 2013. CLAAS-3 slightly underestimates the cloud fraction and reports significantly less ice clouds compared to CALIOP. These findings are expected since CALIOP is more sensitive to thin ice clouds.

When reclassifying optically thin CALIOP data (COT between 0.0 and 0.2) to clear-sky (dashed line in upper panel of Figure 6-2) an overestimation of cloud fraction becomes pronounced. In Figure 6-3 (left panel) the sensitivity of the cloud detection scores is shown as a function of the selected COT threshold. CLAAS-3 is indeed less sensitive to optically thin clouds as the probability of detection (POD) increases with the COT threshold used to distinguish clear and cloudy CALIOP measurements.

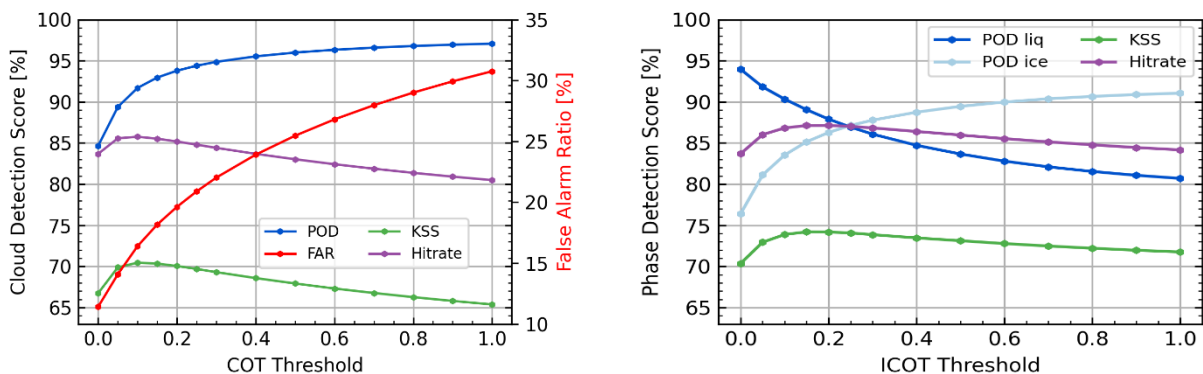


Figure 6-3: Left: CLAAS-3 cloud scores as a function of the COT threshold used to discriminate clear and cloudy CALIOP observations. KSS denotes the Hanssen-Kuipers Skill Score. Right: CLAAS-3 phase scores as a function of the ICOT threshold, which determines the reference CALIOP cloud layer.

However, this does not imply that optically clouds thinner than, say COT=1.0, can generally not be detected by CLAAS-3, because the false alarm ratio also increases with the COT threshold. One could rather say that it is more likely to miss a cloud with CLAAS-3, if it is optically thin. Thus, there are two effects happening simultaneously when increasing the COT threshold (left panel of Figure 6-3):

- Optically thin CALIOP clouds not detected by CLAAS-3 are reset to cloud-free, hence the cloud POD increases.
- Optically thin CALIOP clouds detected by CLAAS-3 are reset to cloud-free, leading to an increased False Alarm Ratio.

The coupling of these effects causes the hit rate and KSS to peak already at COT = 0.1. It means CLAAS-3 systematically misses clouds which are optically thinner than approximately 0.1.

When selecting the CALIOP phase after removing thin cloud layers at the top (dashed line in lower panel of Figure 6-2) the bias compared to CALIOP ice cloud fraction notably reduces but remains negative.

The influence of the choice of the threshold in this analysis is shown in the right panel of Figure 6-3. With increasing threshold the statistics change due to quite some CALIOP profiles turning from ice to liquid phase when comparing cloud-top to lower levels where the threshold is exceeded. This leads to

- Increased liquid POD if the thin ice cloud was incorrectly reported as liquid by CLAAS-3,
- Decreased the liquid POD if the thin ice cloud was correctly reported as ice by CLAAS-3.

Liquid POD, shown by the dark blue line in Figure 6-3, decreases with increasing ICOT threshold, suggesting that the case b) likely prevails in the analysed data.

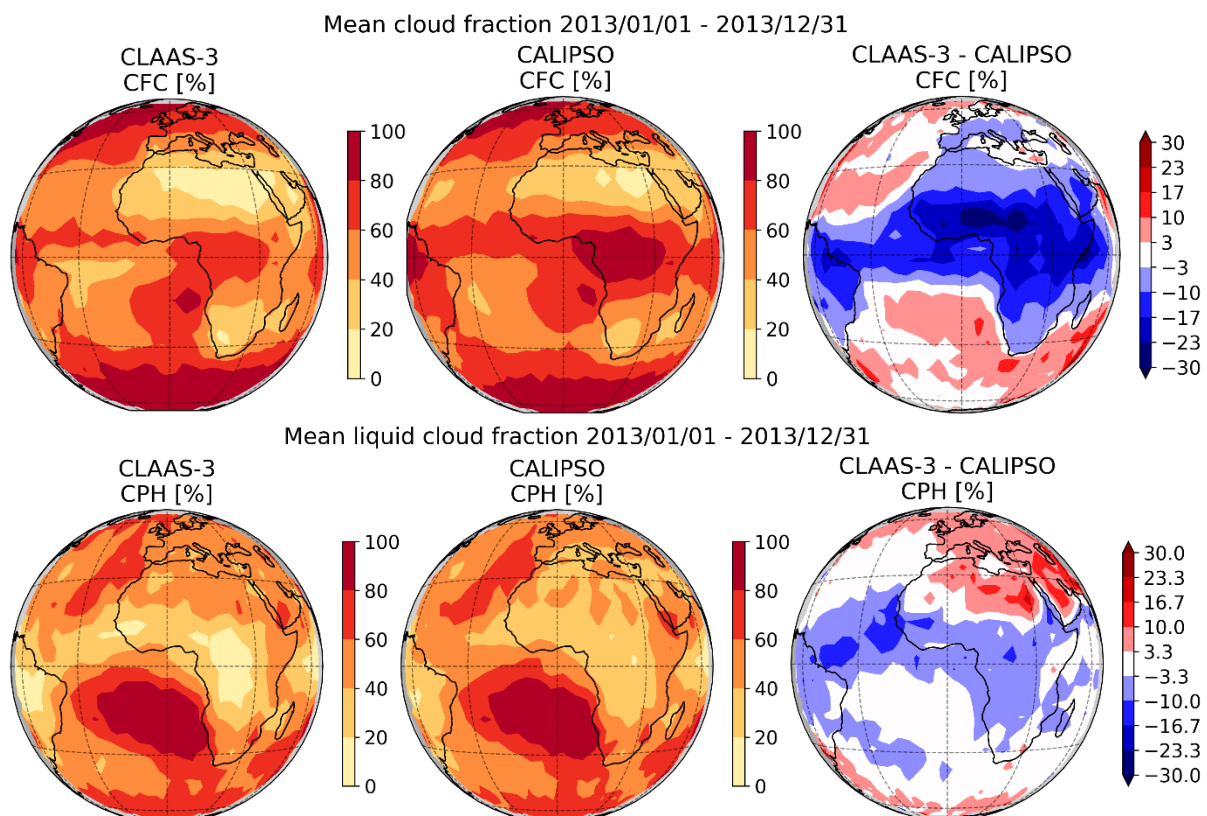


Figure 6-4: Spatial distribution of total (upper row) and liquid (bottom row) cloud fraction as seen by CLAAS-3 (left) and CALIOP (middle), and their difference (right). CALIOP cloud criterion is total column COT > 0. CALIOP liquid phase is taken from the layer where ICOT exceeds 0.2.

In order to investigate the spatial characteristics, all collected cloud mask and phase matchups were remapped to a regular $3.5^\circ \times 3.5^\circ$ grid and averaged within each grid box, as shown in Figure 6-4. For cloud detection the scenario with COT threshold of 0.0 is chosen while for cloud phase the scenario for which the uppermost, thin cloud layers are neglected is shown. The large-scale patterns of both cloud fraction and cloud phase are very similar in CLAAS-3 and CALIOP. However, CALIOP reports much higher cloud fraction in the ITCZ, which can be explained by the large percentage of cirrus clouds in this region, which are more likely to be missed by

CLAAS-3. Over the Tropics, especially over Africa and South America, a negative bias of approximately 10-20% cloud fraction is found. A slight overestimation at the edge of the disk is a known feature of the CLAAS-3 cloud mask product.

For cloud phase, CLAAS-3 liquid cloud fraction is higher than CALIOP over the Sahara and the Arabian Peninsula where the cloud fraction, however, is generally, thus the statistics are based on a relatively low number of observations. CLAAS-3 underestimates liquid cloud fraction in areas with predominantly high convective clouds and outflow cirrus: the filtering of thin high ice cloud layers in CALIOP results in an increased CALIOP liquid cloud fraction, thus becoming larger than the CLAAS-3 values.

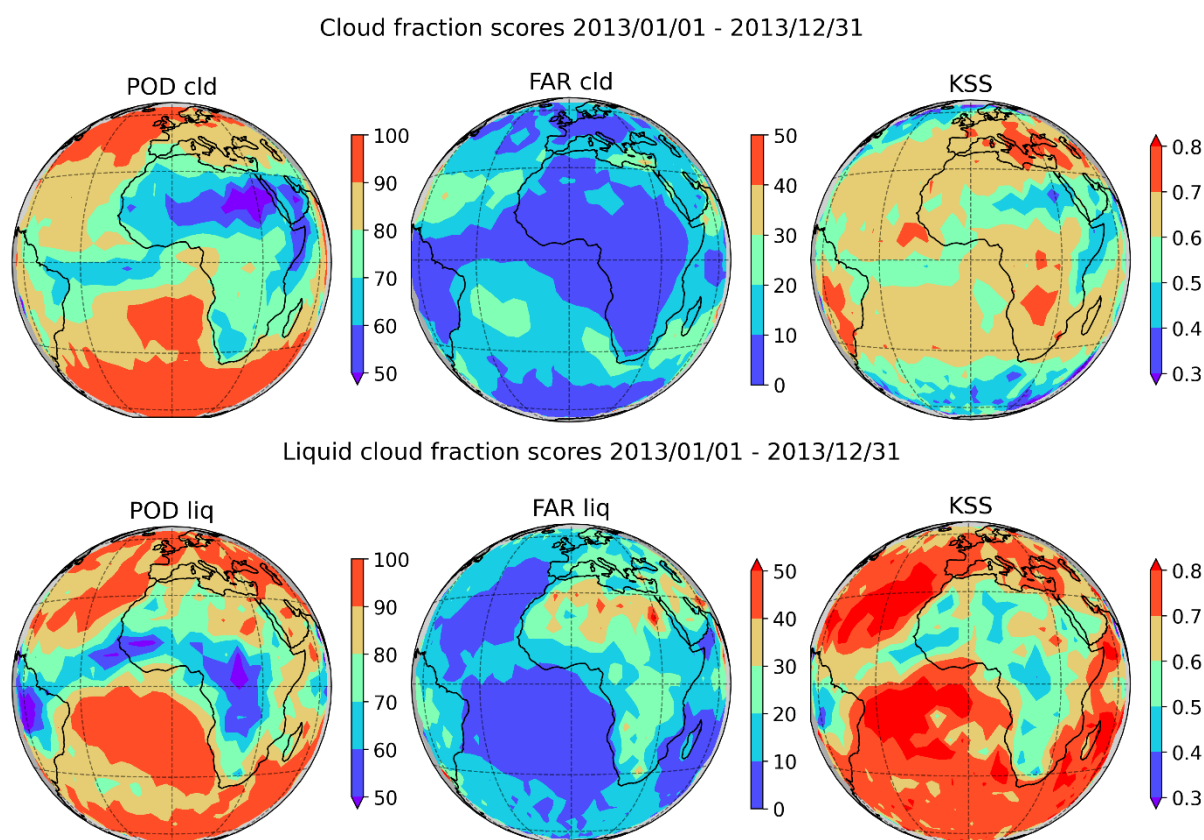


Figure 6-5: Spatial distribution of validation scores of CLAAS-3 against CALIOP: probability of detection (POD), false alarm ratio (FAR) and Hanssen-Kuipers Skill Score (KSS) for total cloud fraction (upper panel) and liquid cloud fraction (bottom panel). CALIOP cloud fraction is taken for COT > 0, and liquid phase is taken from the layer where ICOT exceeds 0.2.

Maps in Figure 6-5 illustrate spatial distributions of cloud detection scores for cloud mask and phase validation. The cloud POD exceeds 90% in large regions, especially over the northern and southern Atlantic. Over the central Atlantic and over land the POD is considerably lower, and the cloud POD minimum lies in the Sahara, where the cloud occurrence frequency is low and the predominating cloud type is very thin cirrus, advected by the outflow from the ITCZ. On the other hand, the cloud FAR is very low, averaging to approximately 10% over the disk. The low cloud POD over the central Atlantic and over Africa is very likely caused by missing high, optically thin clouds.

The phase of liquid stratocumulus clouds over the southern Atlantic is correctly detected with a very high probability > 90% when applying ICOT > 0.2, as shown in the bottom panel of Figure 6-5. However, liquid clouds in the ITCZ are incorrectly classified as ice in about 50% of the cases, but note that liquid cloud fractions are very low there. Simultaneously, the liquid cloud FAR has relatively high values in the same areas, which is caused by classifying thin ice clouds over liquid clouds as liquid.

Looking at the Hanssen-Kuipers Skill Score in the right column of Figure 6-5, a good agreement with spatial patterns of POD can be found. Figure 6-6 illustrates the averaged zonal distribution of KSS for cloud mask and cloud phase. The zonal variability of CPH KSS is quite low between 65°S and 65°N. The cloud mask KSS strongly decreases at high latitudes because of a known tendency of CLAAS-3 to overestimate cloud cover at large viewing angles. Higher KSS values in the tropics with excluded optically thin clouds indicate again that CLAAS-3 misses these clouds. As one can see in the right bottom panel in Figure 6-5 the cloud phase KSS is low over the African continent and the central Atlantic but it is very high over the ocean; due to the averaging the corresponding zonal values are very stable. A slight decrease can be found only in the tropics if excluding clouds with COT < 0.2, as shown in Figure 6-5. Averaged validation scores for cloud mask and cloud phase are shown in Table 6-2.

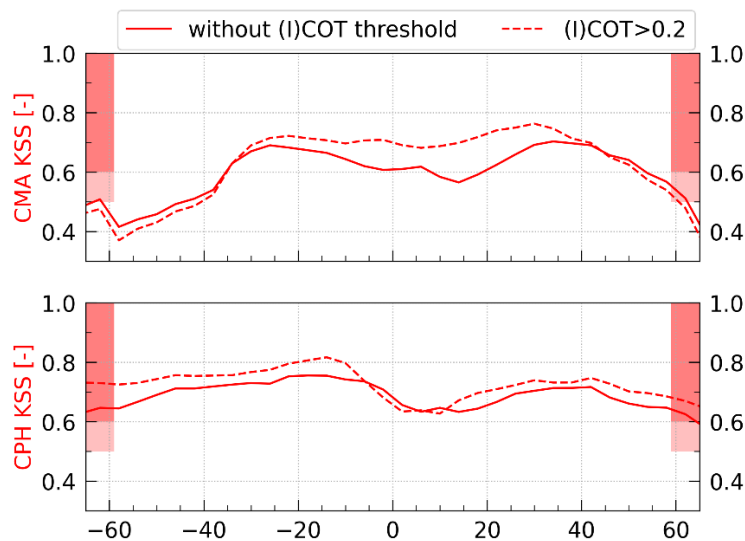


Figure 6-6: CLAAS-3 cloud mask and phase Hanssen-Kuipers Skill score (KSS) in 4-degree latitude bands. Darker and lighter red shadings indicate the target and threshold KSS requirements, respectively. Top: Solid/dashed lines were calculated using total column COT > 0/0.2 as cloud criterion. Bottom: Solid lines refer to the uppermost CALIOP cloud layer, dashed lines refer to the CALIOP cloud layer where ICOT exceeds the 0.2 threshold.

Table 6-2: Summary of Hanssen-Kuipers Skill Scores for CLAAS-3 cloud mask and cloud phase. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.

Parameter & score	CALIOP (I)COT > 0.0	CALIOP (I)COT > 0.2
Cloud fraction KSS [-]	0.67	0.70

Cloud phase KSS [-]	0.70	0.74
---------------------	------	------

6.1.2 Cloud Top Level

CLAAS-3 cloud top pressure, height and temperature are directly comparable to the pressure, height and temperature of the uppermost cloud layer detected by CALIOP. For sensitivity analysis, we also include comparisons against the CALIOP for which the CALIOP reference CTH/CTP/CTT was not taken at the cloud top but rather from a level of the CALIOP profile at which the cloud layer COT (top down) exceeded a certain threshold, i.e. 0.2. This is similar to the approach used for cloud phase validation in Section 6.1.1.

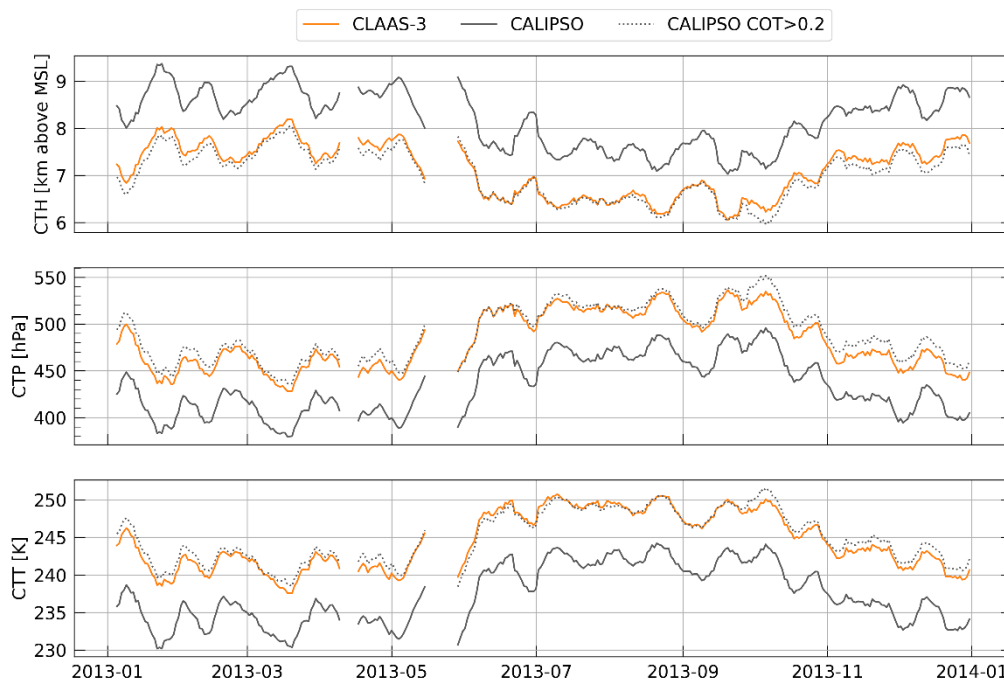


Figure 6-7: Ten-day moving average of CLAAS-3 and CALIOP cloud top products averaged over the full disk. Solid black line denotes the uppermost CALIOP cloud layer; dashed black line represents the CALIOP cloud layer where ICOT exceeds the 0.2 threshold.

A time series of cloud top products from CLAAS-3 and CALIOP in 2013 is shown in Figure 6-7. When comparing against the uppermost CALIOP layer (solid black line), CLAAS-3 significantly underestimates cloud top height (CTH) and consequently overestimates cloud top temperature (CTT) and cloud top pressure (CTP). If CALIOP uppermost cloud layers with ICOT < 0.2 are excluded, the difference to CLAAS-3 is significantly reduced, and the averaged biases of both CTH and CTP fulfil even the optimum requirements. The globally averaged cloud top products from CLAAS-3 and CALIOP almost exactly match in the months June to September when filtered for ICOT < 0.2.

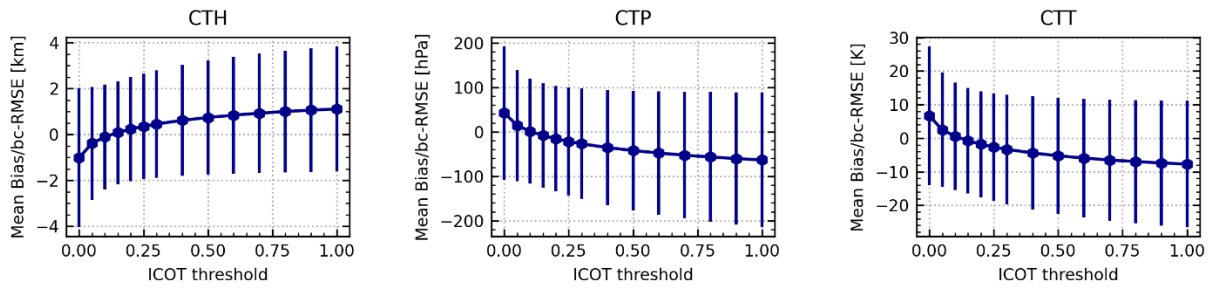


Figure 6-8: Mean cloud top bias (CLAAS-3 – CALIOP) compared against the CALIOP cloud layer where ICOT exceeds a certain threshold. Error bars represent the bias-corrected root mean square error.

In addition to the fixed ICOT threshold of 0.2, we analysed the impact of excluding CALIOP cloud layers as function of varying ICOT threshold. Figure 6-8 shows the biases and corresponding bc-rmse values as error bars for ICOT thresholds between 0.1 and 1.0 as well as the scores for the validation against the uppermost layer (ICOT = 0). Excluding the high thin CALIOP layers strongly influences the bias. The underestimation of CTH and overestimation of CTP/CTT at ICOT = 0 turn into an overestimation of CTH and underestimation of CTP/CTT at ICOT=1.0. The bc-rmse values are smallest if applying an ICOT threshold between 0.1 and 0.3. The best accuracy is achieved with a threshold of ≈ 0.1 . It confirms that CLAAS-3 systematically misses clouds with optical thickness less than 0.1, shown in Section 6.1.1. The overall mean bias and bc-rmse are summarized in Table 6-3.

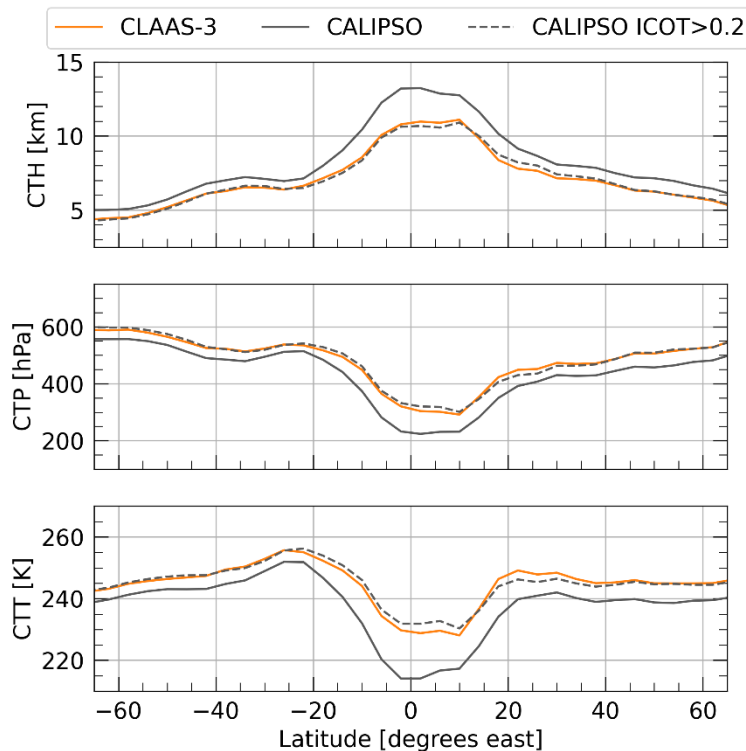


Figure 6-9: Zonal averages of CLAAS-3 and CALIOP cloud top products in 4-degree latitude bands. Solid black line denotes the uppermost CALIOP cloud layer; dashed black line represents the CALIOP cloud layer where ICOT exceeds the 0.2 threshold.

The zonal characteristics of cloud top products from CLAAS-3 and CALIOP are shown in Figure 6-9. The underestimation of CTH and overestimation of CTP/CTT, if comparing with the uppermost CALIOP cloud layer, described above is largest in the tropics and decreases towards high latitudes. This behaviour is expected because optical thin clouds that prevail in that region are mostly missed by CLAAS-3. As seen in the time series, Figure 6-7, the bias is significantly reduced by excluding CALIOP layers containing such clouds. This is the case for cloud top products in all zonal bands.

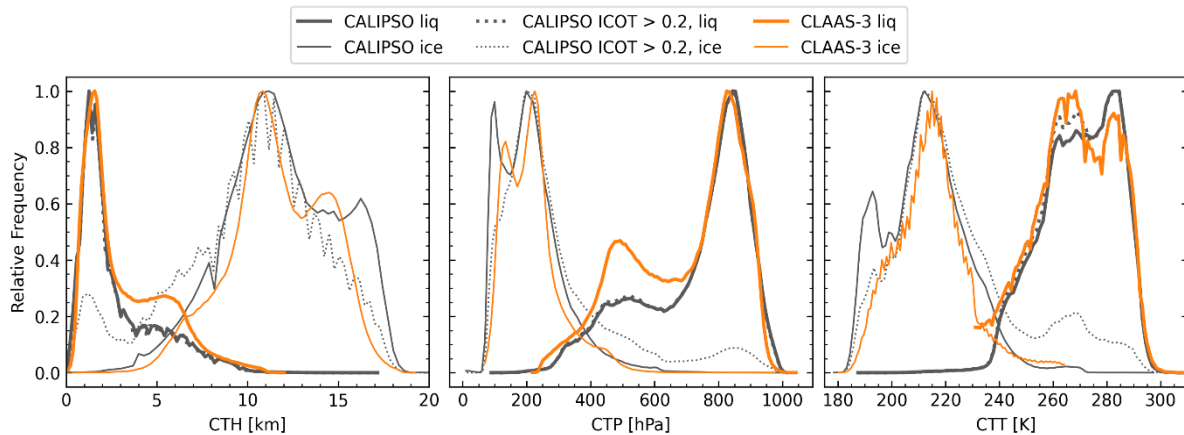


Figure 6-10: 1D-histograms of cloud top products from CLAAS-3 and CALIOP. Solid lines denote liquid phase, dotted lines represent ice clouds.

Next, we compare the relation of cloud top parameters and cloud phase using frequency histograms. The results are shown in Figure 6-10. There is a good agreement between distributions of all parameters for the liquid phase. Note that the difference between the histograms for CALIOP with ICOT > 0.2 and its uppermost layer is small for liquid warm low clouds. Looking at the higher and colder clouds this difference and the difference to CLAAS-3 increases. CLAAS-3 has more liquid clouds at approximately 5km/500hPa, where CALIOP ice clouds with ICOT > 0.2 show higher values. Therefore, CLAAS-3 incorrectly classifies a small amount of the ice mid-level clouds as liquid.

The distributions of ice cloud top products differ more. Both data sets show a bimodal distribution, but for CALIOP it is more pronounced. The second peak in CALIOP ice clouds is related to very high cold clouds, which are also optically thin as seen if comparing the uppermost layer and clouds from the layer where ICOT exceeds 0.2. CLAAS-3 cloud top products contain more cold and high ice clouds than CALIOP with ICOT > 0.2.

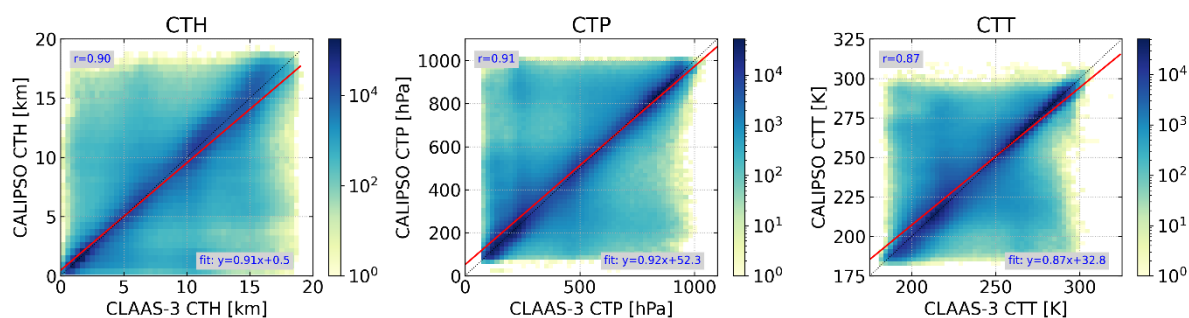


Figure 6-11: Correlation of cloud top products between CLAAS-3 and CALIOP. A dotted line marks the diagonal; red lines show the result of a least squares linear fit. The textbox in the upper left corner displays the Pearson correlation coefficient, the text in the lower right corner the linear fit function. CALIOP values are taken from the layer where ICOT exceeds 0.2.

The 2d-histograms in Figure 6-11 of cloud top parameters from CLAAS-3 and the CALIOP cloud layer with ICOT > 0.2 show the averaged distribution and the correlation of the investigated parameters. Most of the matchups are close to the diagonal line – it confirms a good agreement between cloud top parameters. The overall correlation between cloud top products is quite strong with Pearson correlation coefficients between 0.87 and 0.91. However, high-level cloud top height is slightly overestimated with this ICOT filtering.

Finally, all collected cloud top matchups were remapped to a regular $3.5^\circ \times 3.5^\circ$ grid and averaged within each grid cell in order to compare the spatial features of both datasets (Figure 6-12). The spatial patterns match excellently. The agreement is very good, especially in the northern and southern Atlantic. The cloud top height in these regions does not exceed 200 m and cloud top pressure bias lies under 30 hPa. CALIOP only reports higher clouds, with lower CTP and CTT, over the desert regions in Africa. Since the maps present the CALIOP layers where ICOT exceeds 0.2, the bias in the ITCZ is positive for CTH and negative for CTP and CTT. As it was shown in Figure 6-10, CLAAS-3 has more ice clouds than CALIOP with excluded very thin cloud layers.

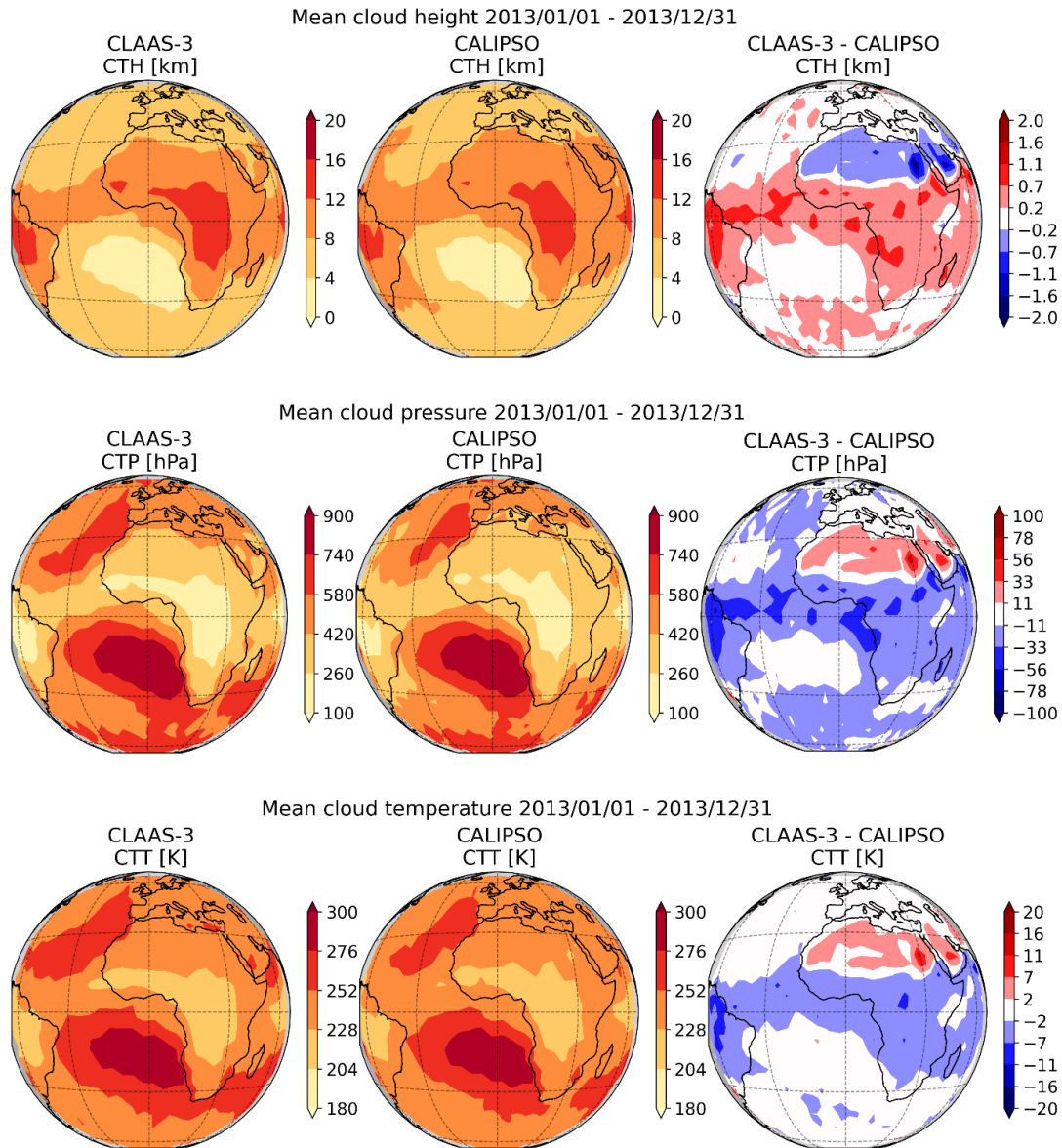



Figure 6-12: Spatial distribution of cloud top products from CLAAS-3 and CALIOP, and their difference. CALIOP values are taken from the layer where ICOT exceeds 0.2.

Table 6-3: Overall validation results for CLAAS-3 cloud top products. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.

Parameter / Score	CALIOP ICOT > 0.0		CALIOP ICOT > 0.2	
	Bias	bc-rmsd	Bias	bc-rmsd
CTH [m]	-1015	3032	230	2260
CTP [hPa]	42	150	-15	119

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

6.1.3 Summary of validation with CALIOP

The results of the validation of CLAAS-3 cloud mask, phase and top products against CALIOP can be summarized as follows:

- CLAAS-3 L2 cloud parameters are generally in better agreement with CALIOP if optically thin clouds are excluded from the latter data set, although the bias in cloud fraction is smallest if all CALIOP clouds are included.
- CLAAS-3 cloud mask and phase fulfil the target requirement, irrespective whether CALIOP optically thin cloud layers are filtered out.
- For the CLAAS-3 cloud top pressure and height, the biases meet the optimum requirement and bc-rmsd meet the target requirement when comparing to CALIOP with ICOT threshold of 0.2.
- The most prominent spatial differences between CLAAS-3 and CALIOP were found in the ITCZ and over the desert regions.
- The validation results are summarized in Table 6-2 and Table 6-3.

6.2 Validation with DARDAR

DARDAR products for all days of March 2013 were collected. DARDAR data were examined in series of 5 profiles. These profiles were, apart from clear-sky, required to consist of (only) the classes ‘ice clouds’, ‘spherical or 2D ice’, ‘supercooled+ice’, ‘highly concentrated ice’, or ‘top of convective towers’. The ice cloud layer properties in the 3 central (from 5) profiles were then integrated vertically and averaged to yield COT, CRE and IWP at ~3 km spatial scale. For CRE a vertical weighting toward the top of the cloud was applied in order to reflect the nature of passive imager CRE retrievals being sensitive to the effective radius in the highest portion of the cloud (Platnick, 2000). The following weighting was applied to obtain the weighted CRE, r_e^{top} :

$$r_e^{\text{top}} = \frac{\int_0^{\text{TOA}} r_e(z) \alpha(z) e^{-\tau(z)/\tau_w} dz}{\int_0^{\text{TOA}} \alpha(z) e^{-\tau(z)/\tau_w} dz}$$

$$\tau(z) = \int_z^{\text{TOA}} \alpha(z') dz'$$

where TOA is top-of-atmosphere, $r_e(z)$ the effective radius as a function of height, $\alpha(z)$ the extinction profile, $\tau(z)$ the integrated cloud optical thickness above height z (so that the total COT is $\tau(0)$), and τ_w the optical thickness determining how far into the cloud the weighting is applied. Note that very large τ_w corresponds to vertically uniform weighting.

The SEVIRI pixel closest to the central DARDAR profile was determined. Based on the time of the DARDAR measurement, the nearest SEVIRI time slot was determined and the corresponding CLAAS retrievals in a 3x3 pixel block were collected. The full 3x3 CLAAS block

was required to be retrieved as ice cloud. CLAAS COT, CRE and IWP were then taken from the centre pixel. For cleaner comparisons, only common collocations among the CLAAS-2.1 and CLAAS-3 datasets were kept.

The collocation procedure described above includes some requirements to guarantee relatively homogeneous scenes, in particular avoiding cloud edges and minimizing potential mismatches due to the parallax effect, which was not corrected for.

Results of the ice cloud optical thickness comparisons are shown in Figure 6-13; for these comparisons DARDAR v3.00 was taken. Logarithmic axes are used because of the large range of COT values. In this log-space a reasonable correlation coefficient of 0.58 between CLAAS-3 and DARDAR ice COT is found. Differences between the two CLAAS-3 retrievals are minor, which is as expected because they employ the same visible channel at 0.6 μm . Differences between CLAAS-2.1 and CLAAS-3 are also minor, except for the maximum retrieved COT, which was 100 in the former and is 150 in the latter data record. Compared to DARDAR, CLAAS-3 has less thin clouds with COT < 1.5, and more clouds with COT > 1.5. As a result, CLAAS-3 has a positive bias of around 2.2.

For ice cloud effective radius the agreement between CLAAS-3 and DARDAR is poor (Figure 6-14). Correlation coefficients are lower than 0.25 and considerable biases of -12 to -22 μm appear. The 1.6 μm retrievals are closer to DARDAR than the 3.9 μm retrievals. To obtain CRE from DARDAR, the profile was weighted to the uppermost part of the cloud with $\tau_w = 1$. If a uniform weighting is applied, the mean DARDAR CRE increases by 4 μm (not shown), and CLAAS-DARDAR biases become correspondingly larger. Note that the CLAAS-3 1.6 mm CRE has increased considerably compared to CLAAS-2.1, which is to a large extent caused by the introduction of an ice cloud model with severely roughened aggregated solid columns (Yang et al., 2013; Baum et al., 2011) replacing the CLAAS-2.1 model based on imperfect hexagonal columns in monodisperse size distributions (Hess et al., 1998).

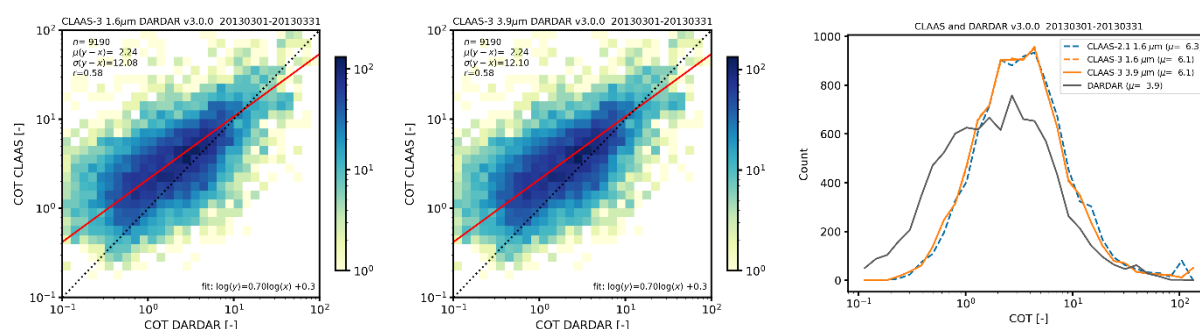


Figure 6-13: Comparison between CLAAS and DARDAR v3.00 ice cloud optical thickness: scatter density plots of CLAAS-3 1.6- μm (left) and CLAAS-3 3.9- μm (centre) ice COT versus DARDAR, and 1D-histograms of CLAAS-3, CLAAS-2.1 and DARDAR ice COT (right). Statistics of the inter-comparison are included in the plots: n is the number of collocations, μ is the mean, σ is the standard deviation or bc-rmsd, and r is the linear Pearson correlation coefficient. In the scatter density plots linear (in log-space) orthogonal fit lines are also included. Note that the two CLAAS-3 COT flavors are nearly identical and their histograms cannot be distinguished. See text for details regarding the collocation procedure.

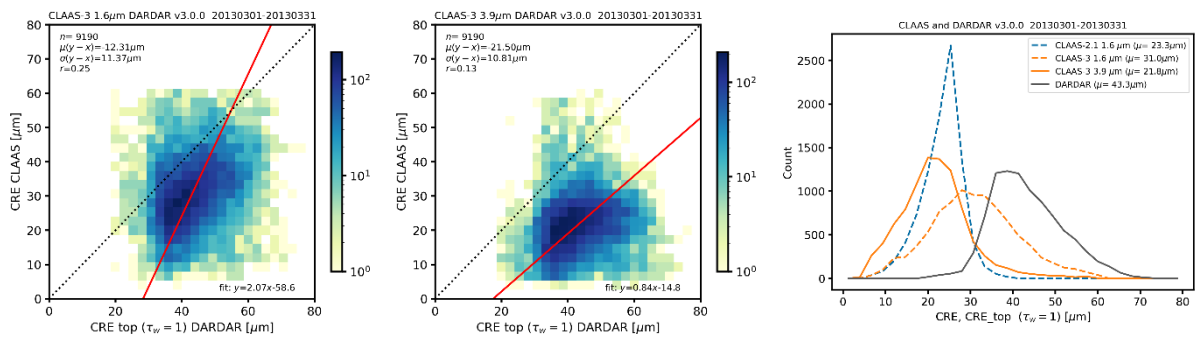


Figure 6-14: As Figure 6-13 but for the ice cloud particle effective radius. For DARDAR a CRE weighted near the cloud top ($\tau_w = 1$) was derived, as explained in the text.

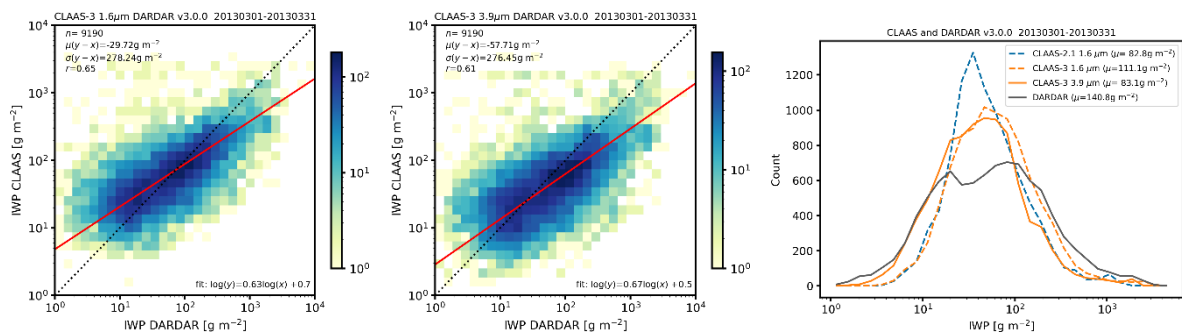


Figure 6-15: As Figure 6-13 but for ice water path.

The ice water path is proportional to the product of COT and CRE, so its validation results resemble the combined results of these cloud properties. The positive bias in CLAAS COT combined with the negative bias in CRE lead to relatively more similar IWP (Figure 6-15). However, there is still a tendency for the CLAAS products to underestimate IWP of thin clouds and overestimate IWP of thick clouds.

Comparable IWP differences have been found before. For example, the distributions in Figure 6-15 look similar to the ones presented in Eliasson et al. (2013), Figure 6. This confirms that an evaluation of passive versus active instruments is difficult due to the different microphysical assumptions (dependent on the spectral region) and the difference between profile information versus column averaged (but weighted to the top of the cloud) measurements. In addition, prior assumptions in the DARDAR effective radius retrievals, in the form of expressions of the particle number concentration as a function of temperature (Cazenave et al., 2019), play a role.

The validation scores are summarized in Table 6-4, both for version 3.00 which was used in the evaluation figures so far, and v.3.10. The difference between these two versions lies in the underlying mass-size relationship, as outlined in Section 11.3. Overall, v3.10 has smaller COT and larger CRE than v3.00. Consequently, the positive CLAAS-3 bias in COT and negative bias in CRE are enlarged. However, the combination of these changes leads to a smaller bias in IWP, being within the optimal and threshold requirement for the CLAAS-3 1.6 µm and 3.9 µm retrieval, respectively. Otherwise the threshold requirements are not met and in particular

the bc-rmsd acquires large values. It is argued that this is not so much due to poor performance of the CLAAS products but rather related to the completely different underlying measurement and retrieval systems.

Table 6-4: Summary of CLAAS IWP validation scores with reference to DARDAR v3.00 and v3.10. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.

DARDAR version	bias (g m ⁻²)		bc-rmsd (g m ⁻²)	
	IWP 1.6 μm	IWP 3.9 μm	IWP 1.6 μm	IWP 3.9 μm
v3.00	-29.7	-57.7	278.2	276.5
v3.10	-1.0	-29.0	235.4	226.7

6.3 Validation with AMSR2

AMSR2 products for all days of March 2013 were collected. For each grid cell in the AMSR2 grid with a valid value in the daytime (ascending node) LWP dataset, a rectangular block of SEVIRI pixels covering that grid cell was sought. Based on the time of the AMSR2 measurement, the nearest SEVIRI time slot was determined and the corresponding CLAAS retrievals were collected. In near-nadir viewing conditions, when the spatial resolution of the SEVIRI is around 3x3 km², this block consists of 9x9 SEVIRI pixels. The CLAAS pixels in the rectangular block were required to be either clear-sky or liquid cloudy, with at least one cloudy pixel. For cleaner comparisons, only common collocations among the CLAAS-2.1 and CLAAS-3 datasets were kept.

First, results are analysed for all collocations (Figure 6-16). With correlation coefficients of 0.86 and 0.79 for CLAAS-3 LWP retrievals based on the 1.6 and 3.9 μm channels, respectively, the overall consistency with AMSR2 turns out to be quite reasonable. Biases are on the order of only a few g m⁻², but there is considerable scatter between the two LWP products, with bc-rmsd between 40 and 50 g m⁻². Both in terms of correlation and bc-rmsd the CLAAS-3 1.6 μm product performs slightly better than the 3.9 μm product. In addition, the CLAAS-3 1.6 μm LWP appears to be a slight improvement compared to the CLAAS-2.1 product. For calculation of the statistics, only the observation pairs inside the plot range have been considered since otherwise a small number (~1%) of CLAAS-3 retrievals with very high LWP have a disproportional impact on these statistics. Note that AMSR2 retrieves negative LWP values, which have been retained in these comparisons.

In terms of spatial distributions both CLAAS-3 retrievals show similar patterns relative to AMSR2. Differences can be broadly summarized as follows. (1) In the stratocumulus areas off the northern and southern African west coast, CLAAS-3 tends to moderately overestimate LWP compared to AMSR2. (2) In the remaining ocean regions at latitudes between 30 °S and 30 °N, frequently dominated by trade cumulus, CLAAS-3 appears to be biased low. This may be partly due to the well-known clear-sky bias in passive microwave-based LWP observations, which persists at low cloud fractions (e.g., Seethala and Horváth, 2010). (3) At higher latitudes,

CLAAS-3 has higher LWP than AMSR2. The latter feature has been observed in earlier studies (e.g., Greenwald, 2009).

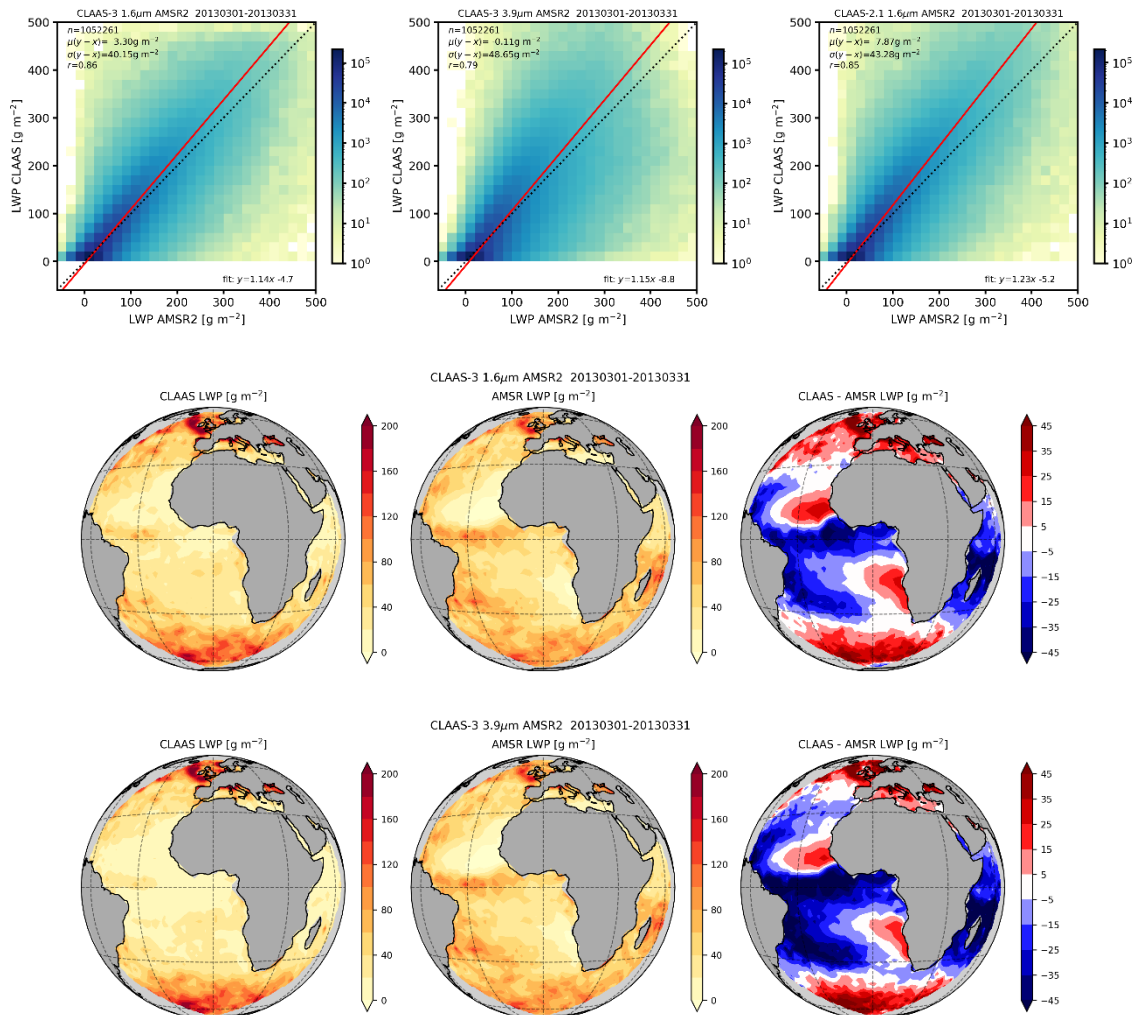


Figure 6-16: Comparison between CLAAS and AMSR2 liquid water path: scatter density plots of CLAAS-3 1.6- μm (top left), CLAAS-3 3.9- μm (top centre) and CLAAS-2.1 1.6- μm (top right) based LWP versus AMSR2, and spatial distribution of CLAAS-3, AMSR2, and their difference for 1.6 μm (middle row) and 3.9 μm (bottom row). Statistics of the inter-comparison are included in the plots: n is the number of collocations, μ is the mean, σ is the standard deviation or bc-rmsd, and r is the linear Pearson correlation coefficient. In the scatter density plots linear orthogonal fit lines are also included. See text for details regarding the collocation procedure.

The AMSR2 retrievals are sensitive to rain, and for LWP larger than about 160 g m^{-2} an empirical separation between cloud and rain water is applied. As a result, these higher values are less reliable, and results are analysed separately for rain-free collocations (Figure 6-17). This changes the statistics, with obviously lower bc-rmsd values, but the findings regarding biases and correlation coefficients remain similar. Note again that, as before, for calculation of the statistics only the observation pairs in the plot range have been considered. The spatial patterns of the datasets and their differences are also qualitatively similar to those resulting from including all collocations.

Passive microwave retrievals are known to be biased in clear-sky conditions (Seethala et al., 2018; Seethala and Horváth, 2010). Therefore, an additional selection criterion was added, namely that all CLAAS pixels corresponding with an AMSR2 grid cell should be cloudy. This reduces the number of collocations by about a factor 2.5 (see Figure 6-18). Perhaps unexpectedly, no improvement in the statistics is observed. The probable explanation for this is that relatively most pixels with low LWP are excluded, which are typically also accompanied by small LWP differences between the datasets. While CLAAS – AMSR2 mean differences are relatively small for the full collocation set, they are considerably larger if only fully cloudy pixels are selected.

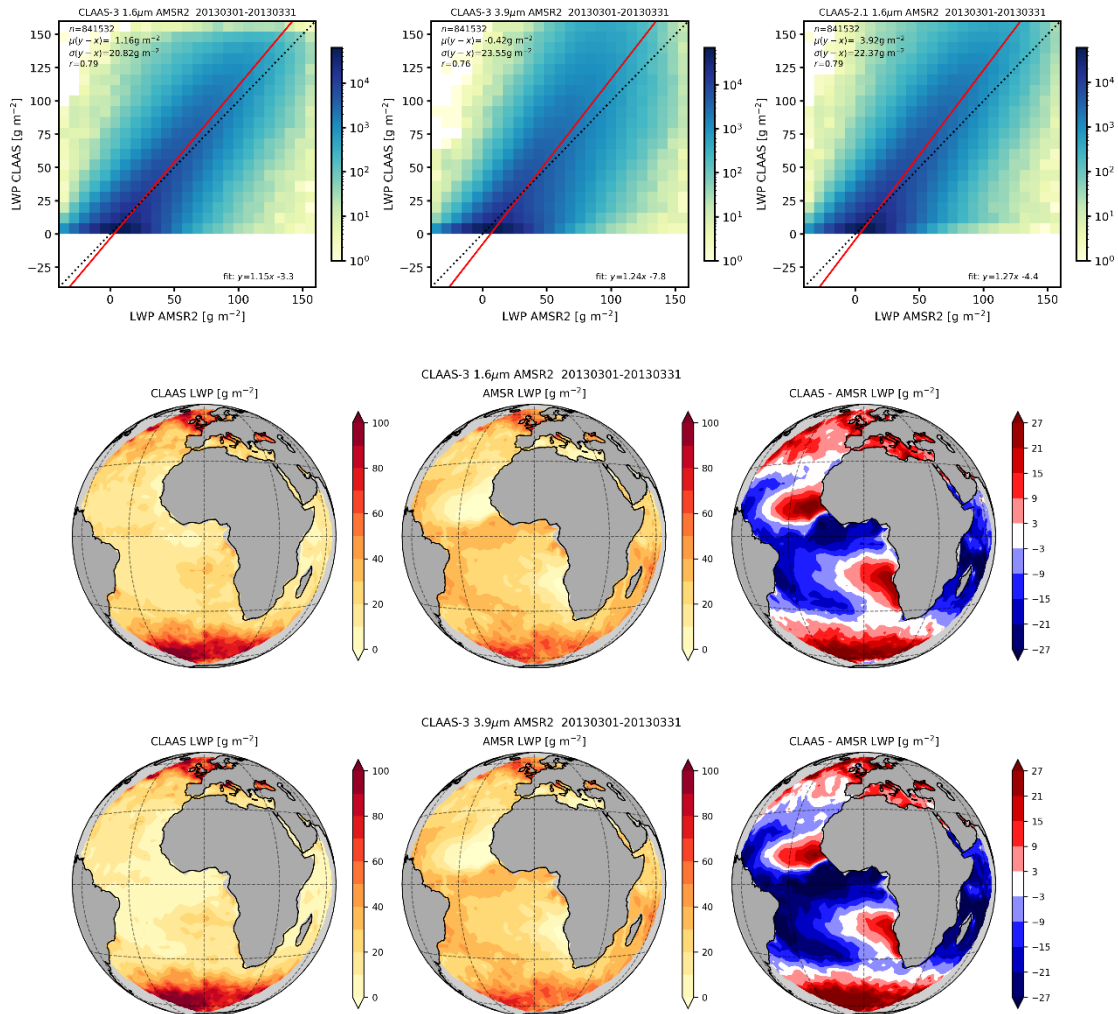


Figure 6-17: Same as in Figure 6-16 but restricted to non-raining pixels.

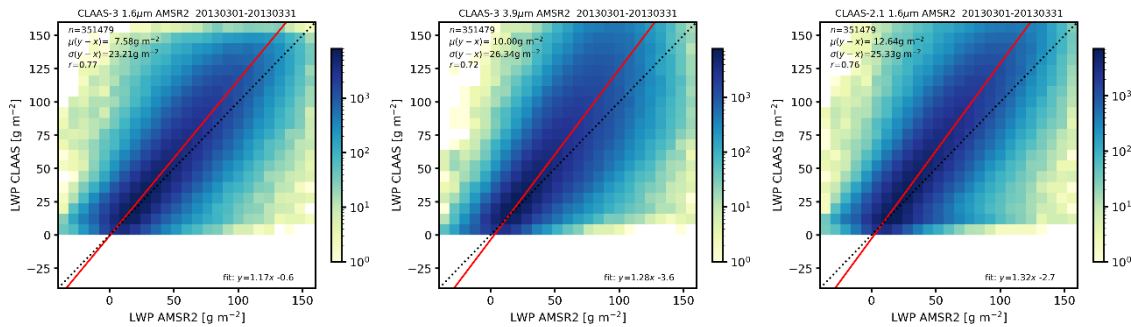


Figure 6-18: Scatter density plots of CLAAS-3 1.6- μm (left), CLAAS-3 3.9- μm (centre) and CLAAS-2.1 1.6- μm (right) based LWP versus AMSR2 for non-raining and fully cloudy pixels.

The collocations with AMSR2 allow evaluating CLAAS error estimates. In Figure 6-19, the error estimates are shown versus the absolute difference between CLAAS and AMSR2 LWP. A clear correlation is found, indicating reasonable skill of the error estimates. On the one hand, the mean difference between the error estimates and the CLAAS-AMSR2 deviations is small, indicating that the error estimates are neither too optimistic nor too pessimistic. On the other hand, one might have expected the error estimates to be smaller than the actual deviations because a number of important error sources, such as deviations from the plane parallel assumption, are not included, and AMSR2 error estimates were not available and thus not considered.

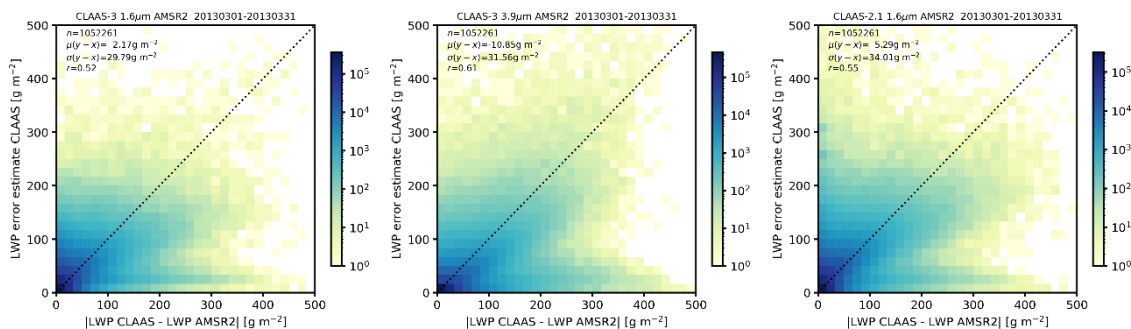


Figure 6-19: CLAAS LWP error estimates versus the absolute difference between CLAAS and AMSR2 LWP: CLAAS-3 1.6 μm (left), CLAAS-3 3.9 μm (centre), and CLAAS-2.1 1.6 μm (right). The collocation dataset is the same as in Figure 6-16.

Validation scores of CLAAS LWP with reference to AMSR2 are summarized in Table 6-5. The scores depend on the selection of collocations but for all three different choices made in that respect, the target requirements are met.

Table 6-5: Summary of CLAAS LWP validation scores with reference to AMSR2. Scores are color-coded: **worse than threshold**, **fulfils threshold**, **fulfils target**, and **fulfils optimal** requirements.

Selection	bias [g m^{-2}]		bc-rmsd [g m^{-2}]	
	LWP 1.6 μm	LWP 3.9 μm	LWP 1.6 μm	LWP 3.9 μm
All collocations	3.30	0.11	40.2	48.7
No rain	1.16	-0.42	20.8	23.6

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

No rain + overcast	7.6	10.0	23.1	26.3
--------------------	-----	------	------	------

6.4 Comparison with MODIS

All MODIS (MOD06 for Terra and MYD06 for Aqua) L2 cloud products within the MSG disk on every 5th day (i.e., 1, 6, 11, 16, 21, 26, and 31) of March 2013 were collected. This selection of days was made to reduce the amount of data. The SEVIRI native grid was downsampled to a 0.2x0.2 degree latitude-longitude grid. Looping over the MODIS granules, for each grid point, a rectangular block of MODIS pixels covering the corresponding SEVIRI pixel was sought. In near-nadir viewing conditions, when the spatial resolution of the SEVIRI and MODIS retrievals is around 3x3 km² and 1x1 km², respectively, this block consists of 3x3 MODIS pixels. If the search was successful, it was furthermore required that all pixels in the MODIS block were cloudy and had a uniform phase. This requirement was extended to a 1-pixel border around the MODIS block to make sure that cloud edges were excluded. Subsequently, the CLAAS observation nearest in space and time to the MODIS observation was selected. If the CLAAS retrieval was cloudy and had the same phase as MODIS, the pair of CLAAS and MODIS observations (the latter averaged over the selected rectangular block) were added to the collocation dataset.

Collocations were performed separately for CLAAS-2.1 and CLAAS-3 and – if applicable – also for 1.6 μm and 3.7/3.9 μm retrievals. In case of Aqua-MODIS, for which retrievals of some cloud parameters are missing due to defect detectors in the 1.6 μm channel, a further requirement, that at least 50% of the MODIS retrievals in the rectangular block were available, was imposed. Unlike for MODIS, the CLAAS CRE retrievals are included in the L2 output even if the observations are located above or below the look-up table space. To properly account for this, collocations with such CLAAS CRE, as well as corresponding CDNC and CGT, retrievals were excluded. This is the same procedure as followed in L3 aggregation. Finally, for cleaner comparisons, only common collocations among the CLAAS-2.1 and CLAAS-3 datasets were kept.

The following subsections describe the inter-comparisons of CTH, COT, CRE for liquid and ice clouds, as well as CDNC and CGT of liquid clouds. To avoid redundancy, LWP and IWP are not separately discussed because these properties are proportional to the product of COT and CRE. However, the statistics of LWP and IWP comparisons are included in the final summary subsection.

The collocations of CLAAS-3 with Terra and Aqua MODIS were inspected separately for the individual instruments and found to be in close agreement. Therefore, all figures and statistics in this section are based on the joint collocation dataset of CLAAS-3 with Terra and Aqua MODIS.

6.4.1 Cloud top height

The comparisons are separated between liquid and ice clouds. Results for liquid cloud top height are shown in Figure 6-20. A reasonable correlation ($r = 0.71$) between CLAAS-3 and MODIS CTH is found but the standard deviation of the differences is quite large ($\sigma = 1175$ m).

From the 1D-histograms it is evident that CLAAS-3 has a narrower distribution with a smaller mode than MODIS. The CLAAS-2.1 1D-histogram, also shown in Figure 6-20, is in comparison to the others rather jerky because CLAAS-2.1 CTH takes only a relatively small number of values, and it also tends toward a bimodal rather than a unimodal distribution.

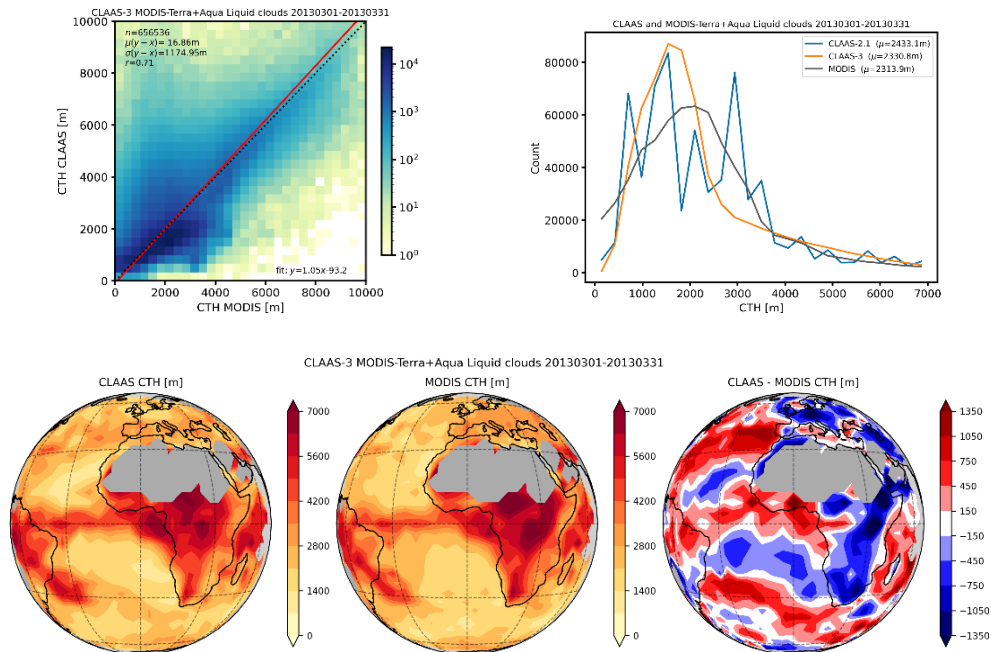


Figure 6-20: Comparison between CLAAS and MODIS cloud top height of liquid clouds: scatter density plot of CLAAS-3 versus MODIS C6.1 (top left), 1D-histograms of CLAAS-3, CLAAS-2.1 and MODIS (top right), and spatial distribution of CLAAS-3, MODIS, and their difference (bottom left to right). Statistics of the inter-comparison are included in the plots: n is the number of collocations, μ is the mean, σ is the standard deviation, and r is the linear Pearson correlation coefficient. In the scatter density plot a linear orthogonal fit line is also included. See text for details regarding the collocation procedure.

For ice clouds (Figure 6-21) a systematic difference between CLAAS-3 and MODIS is apparent: the former yielding almost 1 km higher cloud tops. Moreover, CLAAS-3 CTH has a clearly bimodal distribution with peaks near 10-11 and 14-15 km, while MODIS and also CLAAS-2.1 have a unimodal distribution with a mode around 9 km. The bimodal shape of the CLAAS-3 CTH histogram is also found in CALIPSO measurements (see left panel in Figure 6-10), so CLAAS-3 appears to be more consistent with CALIPSO measurements than MODIS and CLAAS-2.1 in this respect. From the spatial distributions shown in Figure 6-21 it can be concluded that the overall higher CLAAS-3 cloud tops compared to MODIS occur almost uniformly over the MSG disk.

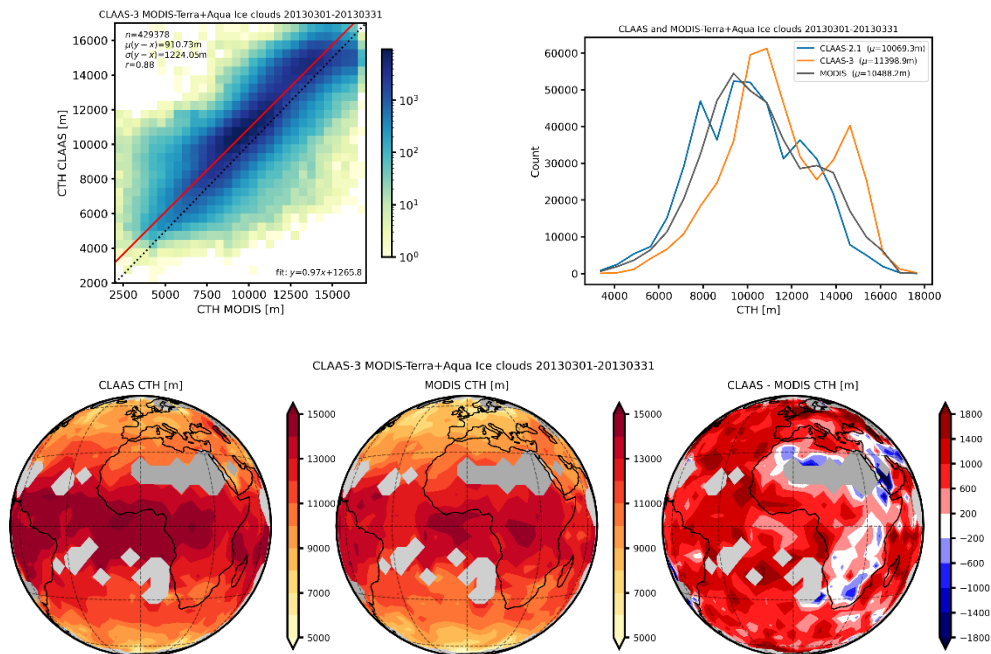


Figure 6-21: As Figure 6-20 but for ice clouds.

6.4.2 Cloud optical thickness

The various data records contain multiple COT retrievals. Corresponding 1D-histograms are shown in the top right panel of Figure 6-22. As confirmed by these histograms, the dependence of the COT retrieval on the SWIR channel is weak and therefore only the 3.7-/3.9- μm based retrievals are shown in the other panels. Overall, CLAAS COT is smaller than MODIS by about 3. If the MODIS pixels in the block collocated with a SEVIRI pixel are logarithmically instead of linearly averaged, which makes most sense for this variable, the difference remains but is reduced to 2. The maps indicate that the CLAAS-3 and MODIS spatial distributions are very consistent, with differences smaller than 1 over most of the disk with exception of the high latitudes. This suggests a dependence on solar zenith angle. Indeed, Figure 6-23 shows that CLAAS and MODIS COT are very similar for SZA smaller than about 50 degrees, whereas for larger SZA the CLAAS-MODIS COT difference becomes increasingly more negative. This seems particularly due to MODIS COT (and LWP) acquiring too large values at high SZA (Greenwald, 2009; Seethala and Horváth, 2010).

For ice clouds no systematic difference between CLAAS-3 and MODIS is found (Figure 6-24). On the other hand, the standard deviation of the difference ($\sigma = 18.6$) is larger than for liquid clouds ($\sigma = 12.0$). A pattern of increasing (negative) differences towards large SZA is present, but far less pronounced than for liquid clouds.

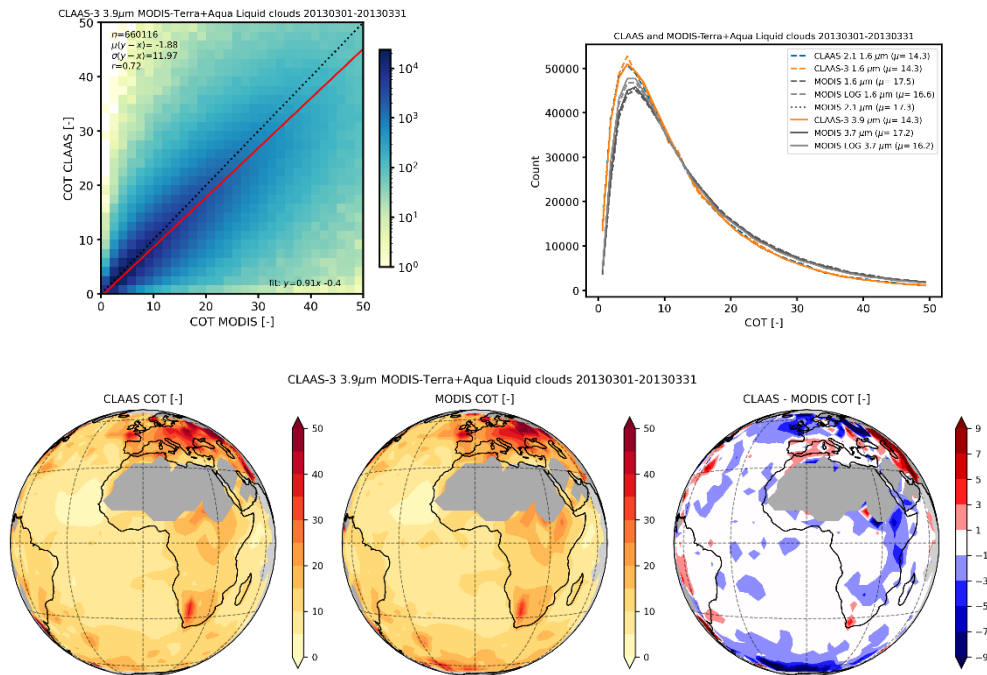


Figure 6-22: As Figure 6-20 but for liquid cloud optical thickness. The scatter density plot (top left) and maps (bottom) are compiled from the 3.9- μm based CLAAS-3 COT and the 3.7- μm based MODIS logarithmically-averaged COT. 1D-histograms (top right) are shown for retrievals based on different SWIR channels as indicated, and for linearly and logarithmically (indicated by 'LOG') averaged MODIS retrievals.

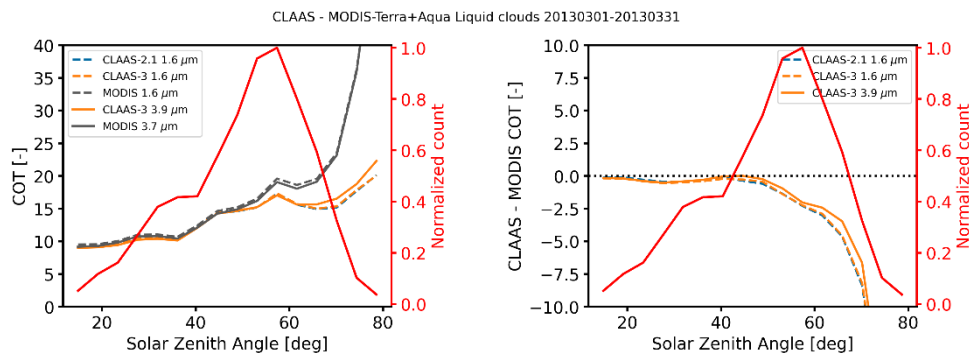


Figure 6-23: Liquid cloud optical thickness as a function of solar zenith angle: for CLAAS and MODIS separately (left) and CLAAS-MODIS difference (right). The normalized number of observations per solar zenith angle bin is indicated by the red line in both panels.

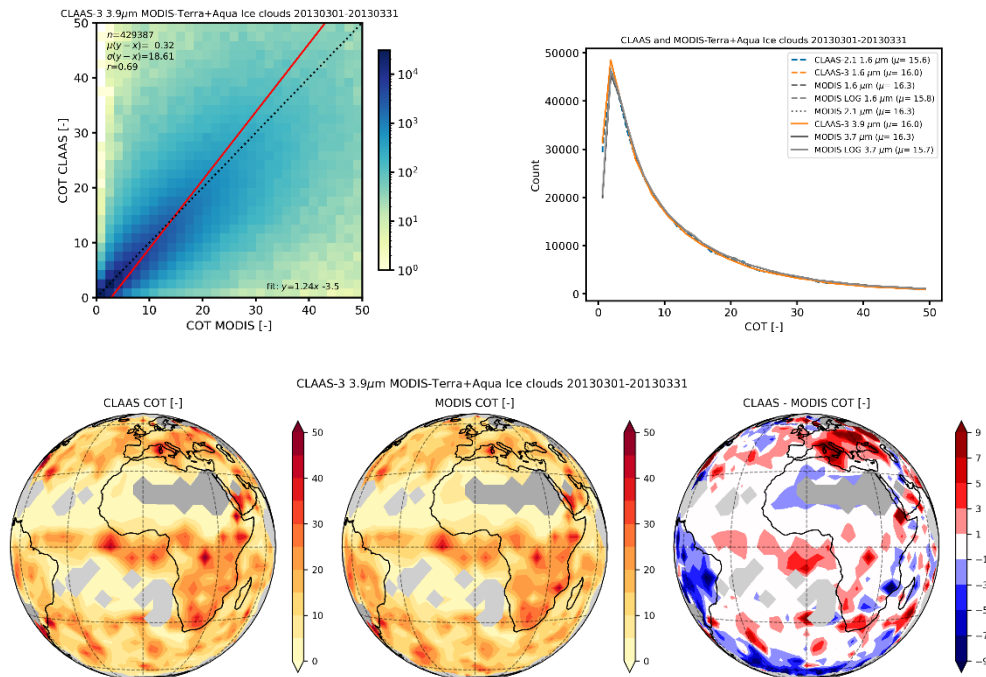


Figure 6-24: As Figure 6-22 but for ice cloud optical thickness.

6.4.3 Cloud particle effective radius

Cloud particle effective radii are markedly different between the 1.6 and 3.9/3.7 µm retrievals. Therefore, Figure 6-25 shows scatter density plots and maps for both these channels. Both for CLAAS-3 and MODIS the 1.6 µm based CRE histograms are broader than those based on 3.9/3.7 µm. The agreement between CLAAS-3 and MODIS is somewhat better for 3.9/3.7 µm than for 1.6 µm, both in terms of bias and standard deviation as well as correlation coefficient. Spatial distributions of the respective datasets are rather similar. For the 1.6 µm CRE retrievals the maps suggest a dependence on SZA, as was observed for COT. Figure 6-26 confirms this, indicating that the CLAAS-MODIS bias is smaller than 1 µm for SZA < 40 degrees, but considerably larger (CLAAS-3 and -2 having smaller CRE than MODIS) toward high SZA. Overall, the CLAAS-3 mean CRE is 0.7 mm smaller than CLAAS-2.1, which can be largely attributed to the choice of a narrower droplet size distribution in the CLAAS-3 liquid cloud model (see also Benas et al., 2019). For the 3.9/3.7 µm CRE retrievals an opposite, and much weaker, dependence of the CLAAS-MODIS bias on SZA is found. In fact, the maps suggest a somewhat stronger dependence of the bias on the SEVIRI viewing angle, which can indeed be confirmed by a corresponding dependency plot (not shown).

When comparing the CLAAS-3 maps and 1D-histograms in Figure 6-25, an apparent inconsistency shows up. From the maps it looks like the 3.9/3.7 µm CRE is much smaller than the 1.6 µm CRE. However, the overall mean CLAAS-3 CRE values, included in the histogram plot, are nearly identical (12.4 µm and 12.6 µm, respectively). The explanation for this is the occurrence of relatively larger 3.9/3.7 µm and smaller 1.6 µm CLAAS-3 CRE at high SEVIRI viewing angles, which are hardly visible in the maps.

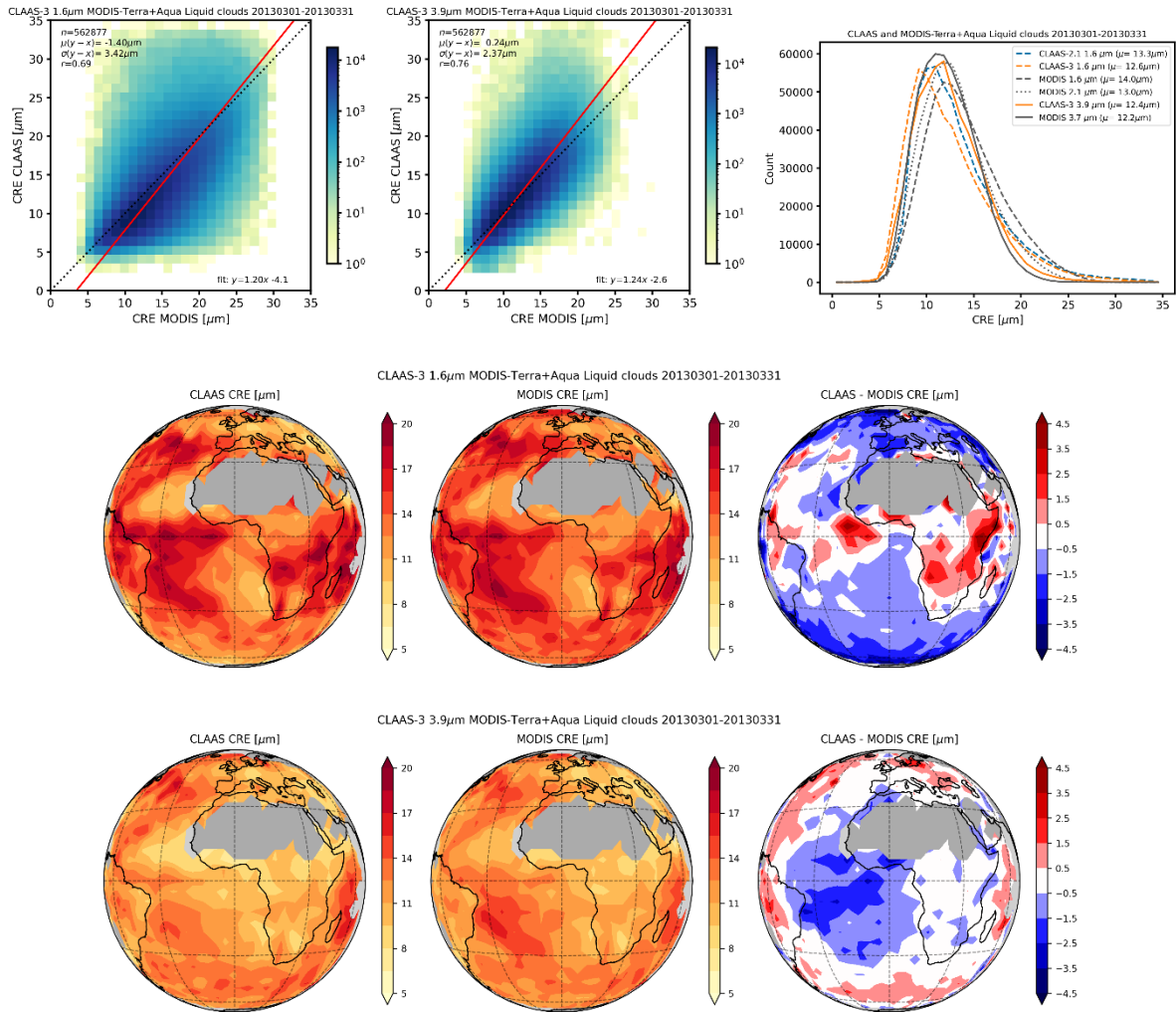


Figure 6-25: Comparison between CLAAS and MODIS liquid cloud droplet effective radius: scatter density plot of CLAAS-3 versus MODIS C6.1 CRE based on 1.6 µm (top left), and on 3.9/3.7 µm (top middle), 1D-histograms of CLAAS-3, CLAAS-2.1 and MODIS CRE (top right), spatial distribution of CLAAS-3, MODIS, and their difference for CRE based on 1.6 µm (middle row), and on 3.9/3.7 µm (bottom row).

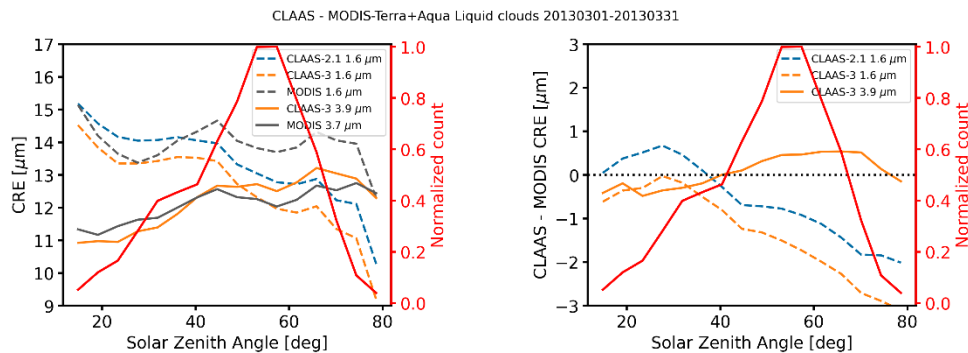


Figure 6-26: Liquid cloud droplet effective radius as a function of solar zenith angle: for CLAAS and MODIS separately (left) and CLAAS-MODIS difference (right). The normalized number of observations per solar zenith angle bin is indicated by the red line in both panels.

Comparisons for ice clouds are shown in Figure 6-27. Here, the histograms for the 1.6 μm retrievals yield very different results, i.e. much larger particles with a wider size range, than those based on 3.9/3.7 μm, both for the CLAAS-3 and MODIS datasets. It is consistent with an effective radius of ice crystals increasing from the top down into the cloud (e.g., Van Diedenhoven et al., 2016), in combination with a larger penetration depth of photons at 1.6 μm. CLAAS-3 CRE is on average smaller than MODIS CRE in both channels, with the largest bias (of -3.5 μm) occurring at 3.9/3.7 μm. The overall relatively good agreement between CLAAS-3 and MODIS ice CRE is a good sign and was also expected since both retrievals rely on the same ice cloud model with severely roughened aggregated solid columns (Yang et al., 2013; Baum et al., 2011). In contrast, CLAAS-2.1 single scattering properties are based on imperfect hexagonal columns in monodisperse size distributions (Hess et al., 1998), yielding much smaller CRE retrieval than the CLAAS-3 1.6 μm retrievals, and coincidentally a quite similar CRE histogram as the CLAAS-3 3.9 μm retrievals.

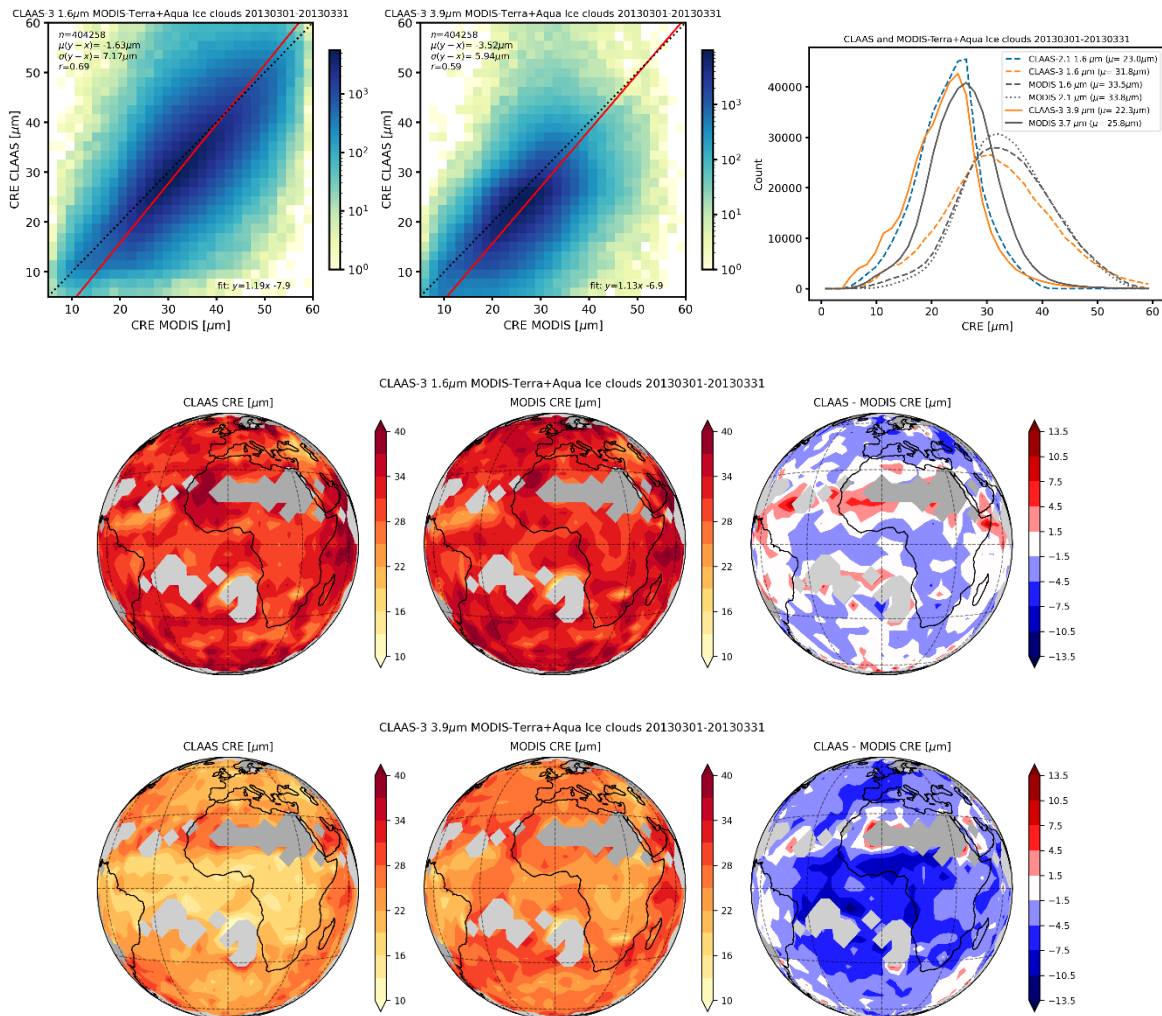


Figure 6-27: As Figure 6-25 but for the ice cloud particle effective radius.

6.4.4 Liquid cloud droplet number concentration and geometrical thickness

Liquid cloud droplet number concentration (CDNC) and cloud geometrical thickness (CGT) are derived from the 3.9/3.7 μm based COT and CRE. In the MODIS L2 files these variables are not included. Therefore, they were derived from the COT and CRE products using the same relations as used in CLAAS-3, i.e. following Bennartz and Rausch (2017). CDNC has a strong inverse ($\sim r_e^{-5/2}$) dependency on CRE. As a result, the negative CRE bias of CLAAS-3 compared to MODIS in the central part of the MSG disk, shown in the bottom panel of Figure 6-25, translates to a positive CDNC bias (Figure 6-28). Towards the edge of the disk the bias is reversed. As indicated by the 1D-histograms the overall agreement is rather good with CLAAS-3 having a smaller mode than MODIS but a nearly identical mean CDNC. The geometrical thickness (CGT) is relatively more dependent on COT than on CRE and agrees reasonably well between CLAAS-3 and MODIS (see Figure 6-29), not showing particular spatial patterns of differences as for CDNC.

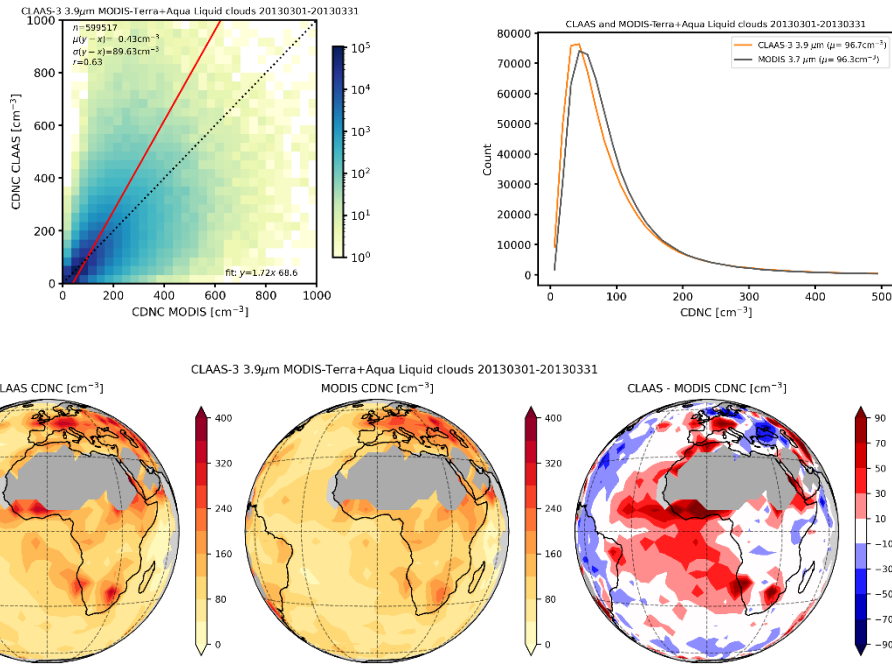


Figure 6-28: Comparison between CLAAS and MODIS liquid cloud droplet number concentration: scatter density plot of CLAAS-3 versus MODIS C6.1 CDNC (top left), 1D-histograms of CLAAS-3 and MODIS CDNC (top right), spatial distribution of CDNC from CLAAS-3, MODIS, and their difference (bottom row).

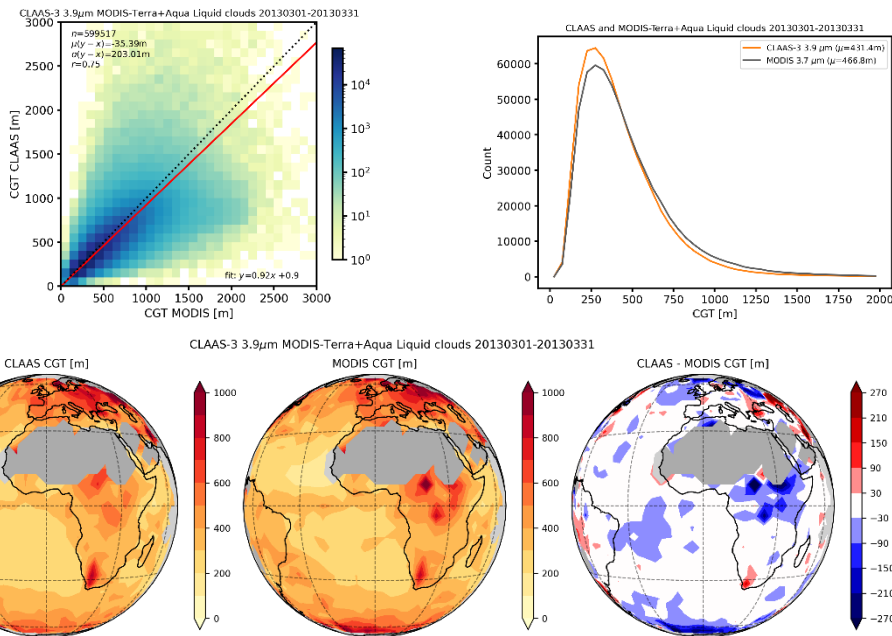


Figure 6-29: As Figure 6-28, but for liquid cloud geometrical thickness.

6.4.5 Summary of comparison with MODIS

In this section the comparisons with MODIS are summarized in terms of the bias and bc-rms differences (Table 6-6). As argued before, these comparisons have been performed to gain

insight in the characteristics of the products, but cannot be considered as validation. Therefore it is not strictly necessary to meet the requirements.

Table 6-6: Summary of CLAAS-MODIS comparison scores: mean difference, bc-rmsd, and correlation coefficient. Where applicable, results for the mean difference and bc-rmsd are color-coded with reference to the L2 requirements: **worse than threshold**, **fulfils threshold**, **fulfils target**, and **fulfils optimal** requirements.

Variable	mean difference	bc-rmsd	correlation coeff.
Liquid CTH	17 m	1175 m	0.71
Ice CTH	911 m	1224 m	0.88
LWP 1.6 μm	-31 g m^{-2}	86 g m^{-2}	0.75
LWP 3.9/3.7 μm	-15 g m^{-2}	84 g m^{-2}	0.74
IWP 1.6 μm	-34 g m^{-2}	200 g m^{-2}	0.76
IWP 3.9/3.7 μm	-36 g m^{-2}	168 g m^{-2}	0.73
Liquid COT 1.6 μm	-2.3	11.9	0.73
Liquid COT 3.9/3.7 μm	-1.9	12.0	0.72
Ice COT 1.6 μm	0.2	18.7	0.68
Ice COT 3.9/3.7 μm	0.3	18.6	0.69
Liquid CRE 1.6 μm	-1.4 μm	3.4 μm	0.69
Liquid CRE 3.9/3.7 μm	0.2 μm	2.4 μm	0.76
Ice CRE 1.6 μm	-1.6 μm	7.2 μm	0.69
Ice CRE 3.9/3.7 μm	-3.5 μm	5.9 μm	0.59
Liquid CDNC	0.4 cm^{-3}	89.6 cm^{-3}	0.63
Liquid CGT	-35 m	203 m	0.75

7 Evaluation of CLAAS-3 aggregated (L3) cloud parameters

This section covers the evaluation of CLAAS-3 L3 products. These consist of daily and monthly aggregations as well as monthly mean diurnal cycles. The evaluation is organized according to Table 7-1.

Table 7-1: Overview of reference datasets used for the evaluation of CLAAS-3 L3 parameters.

Section	Reference observations	Parameters
7.1	SYNOP	CFC
7.2	CALIOP	CFC, CTH
7.3	UWisc	LWP
7.4	MODIS	CFC, CTH, CTP, CPH, LWP, IWP, JCH

7.1 Validation with SYNOP

In this section, the monthly mean CLAAS-3 cloud fractional cover (CFC) is compared against SYNOP data. For these comparisons the CLAAS-3 L3 CFC values were averaged over an area of 5x5 pixels surrounding the SYNOP station in order to reflect the typical spatial extent of CFC observations by human observers. The complete time-span from 2004-2020 is validated.

All available SYNOP reports in the geographical domain covered by the Meteosat full disk field of view were taken into account. Nevertheless, the weather station records from the SYNOP database had to be preselected on the basis of following criteria:

- Only manned airport stations were considered for the validation, as described in Section 11.1.
- Stations which cover more than 95% of the full time period from 2004 to 2020 are included. This balances the increasing total number of stations and decreasing number of manned stations with time and provides a homogeneous time series of observations at each SYNOP site.
- Observations with very high SEVIRI viewing zenith angles (VZA) are omitted and only those stations that are within 75° VZA of SEVIRI are used here.
- Monthly means of SYNOP cloud fraction are used for this comparison. A valid monthly mean at a SYNOP site - the same constraints are applied to CLAAS-3 - is an average of at least 20 daily means, whereas for a daily mean at least 6 measurements had to be found.

The geographical distribution of selected SYNOP stations is unbalanced. As shown in Figure 7-1, the majority of the stations are located in the northern mid-latitudes and tropics while there are fewer stations over Africa and in the southern hemisphere. This uneven distribution has to be kept in mind when looking at accumulated statistics.

Figure 7-1 shows the mean difference between CFC for CLAAS-3 and SYNOP at all selected SYNOP sites averaged for the whole period from 2004 to 2020. An overestimation of CLAAS-3 monthly means of CFC is found in the middle East and Arabia. These regions have a high surface albedo, which may give rise to false cloud detection in CLAAS-3, although this is not found in the comparisons with CALIOP (see Figure 6-4) . The bias over central Europe does not exceed $\pm 10\%$ and hence lies within the approximate accuracy of SYNOP observations. CLAAS-3 CFC monthly means are smaller over the Iberian Peninsula, northern Africa and at the coastal stations in South America. The station observations at the coast are difficult to compare with satellite measurements because the matched SEVIRI pixels can be classified as water pixels whereas the SYNOP observation's location is on land.

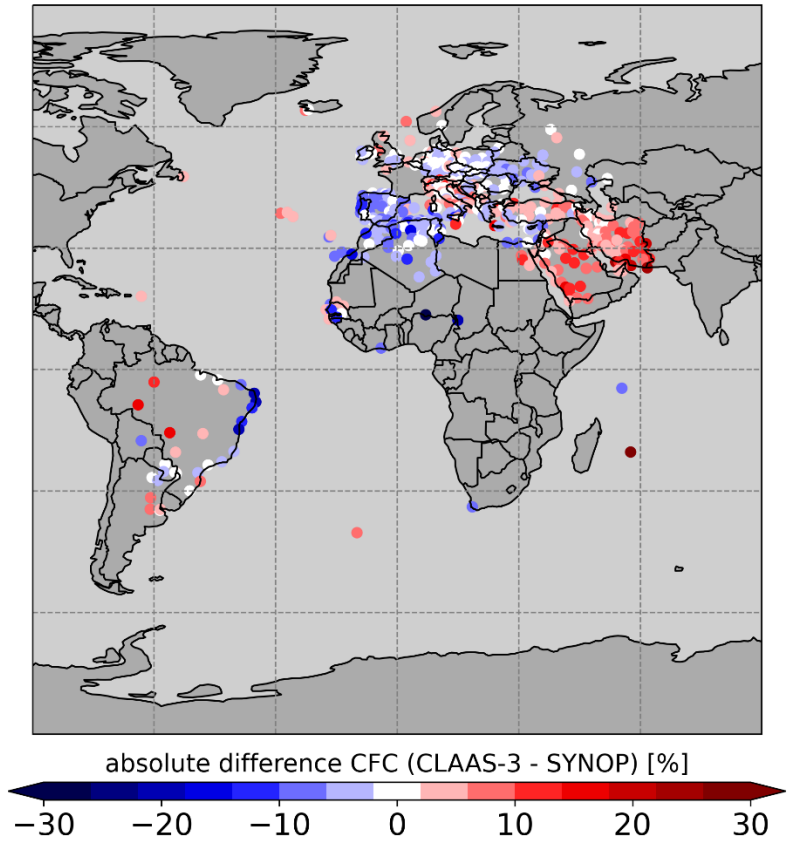


Figure 7-1: Mean difference between CLAAS-3 and SYNOP cloud cover at each preselected SYNOP site for the entire period 2004-2020.

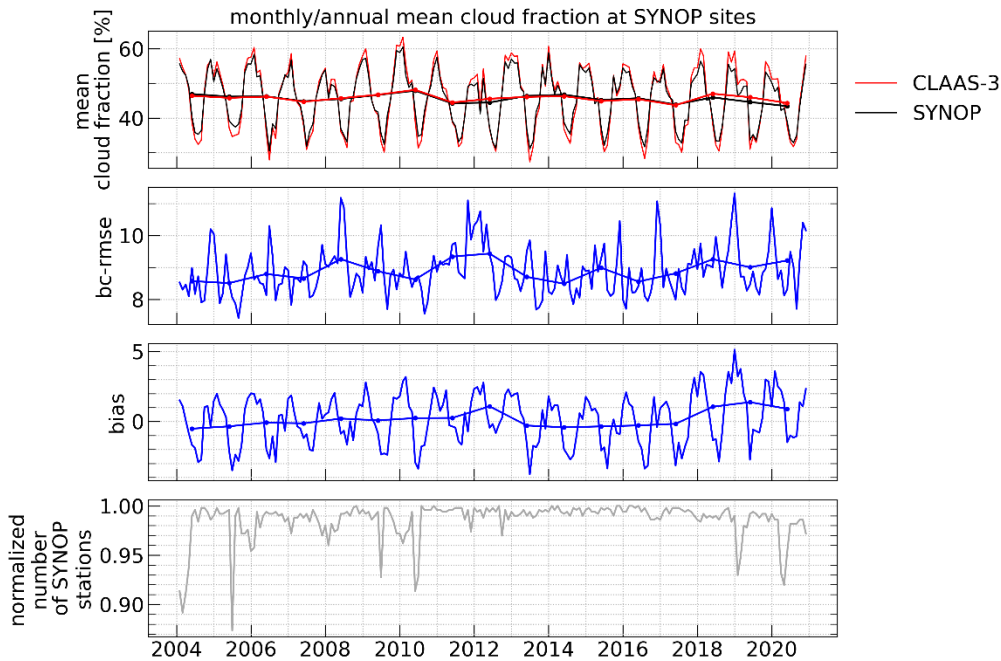


Figure 7-2: Time series of mean cloud cover for CLAAS-3 (red), and SYNOP (black) (upper panel), bias-corrected rmse (second panel), bias (third panel), and the number of stations (lower panel) normalized to 1 for the entire period 2004-2020.

The complete time series of SYNOP and CLAAS-3 monthly means aggregated over the entire domain is shown in Figure 7-2. Both time series show a seasonal cycle with maximum cloud cover in winter and minimum cover during summer. This is due to the fact that the majority of SYNOP stations is located in the northern hemisphere on land, e.g. in Central Europe. So, the course of the time series is mainly determined by the seasonal cycle of cloudiness in the northern hemisphere.

Both time series show a very good overall agreement. The bias, in the third panel, is very stable and lies within $\pm 5\%$ CFC for the entire period. Only after 2018 a slightly higher bias is observed. The bias corrected root mean square error (bc-rmse), shown in the second panel, becomes more stable after 2013, with the switch to MSG-3, and lies generally below 12%. The averaged bias of 0.2% fulfils the optimum requirement, and for the averaged bc-rmse of 8.89% the target requirement is met, see Table 7-2.

As stated earlier, also the SYNOP measurements can be inaccurate with a more small-scale geometry. The so-called scenery effect leads to an overestimation of cloudiness by SYNOP due to the obscuring of cloud-free spaces by convective clouds with high vertical extent. In some regions, for example Scandinavia this effect shows a seasonal cycle (Karlsson, 2003). On the other hand, cloud cover is often underestimated in synoptical measurements at nighttime because of difficulties in observing semi-transparent cirrus clouds. In the monthly mean time-series all these effects are averaged, but no further conclusions can be drawn on that subject within this report. A temporally and spatially higher resolved study, possibly separated for the different cloud types would be needed to investigate the individual effects more closely.

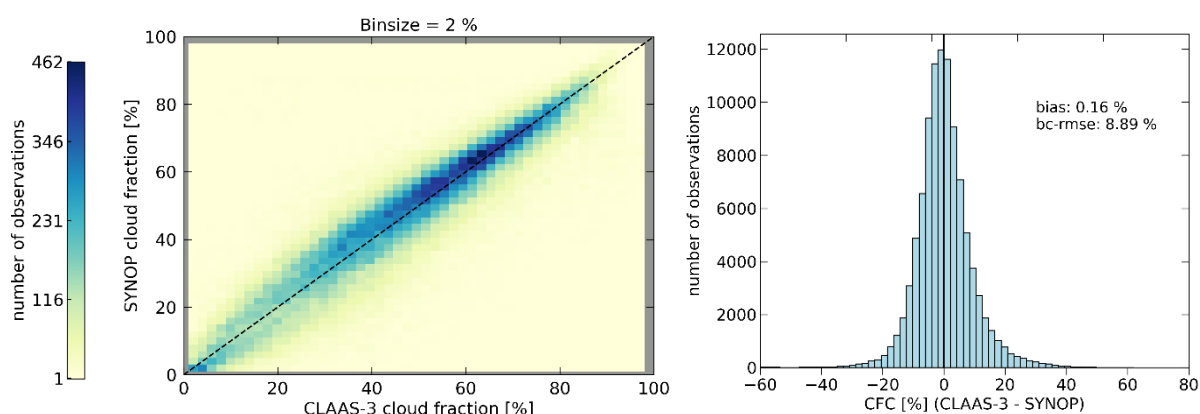


Figure 7-3: Validation of monthly mean CFC from CLAAS-3 with SYNOP: scatter plot (left) and frequency distribution of deviation (right).

More statistical aspects of the validation results are presented in Figure 7-3. The left plot shows the monthly mean CFC values for CLAAS-3 and SYNOP as a scatter plot. Here it becomes clear that the bias is close to zero, that most observations lie between 30 % and 80 % averaged cloud cover, and that for small CFC CLAAS-3 slightly overestimates SYNOP reports. In the right panel, the frequency distribution of the difference CLAAS-3 – SYNOP assures that the differences are generally small and evenly distributed.

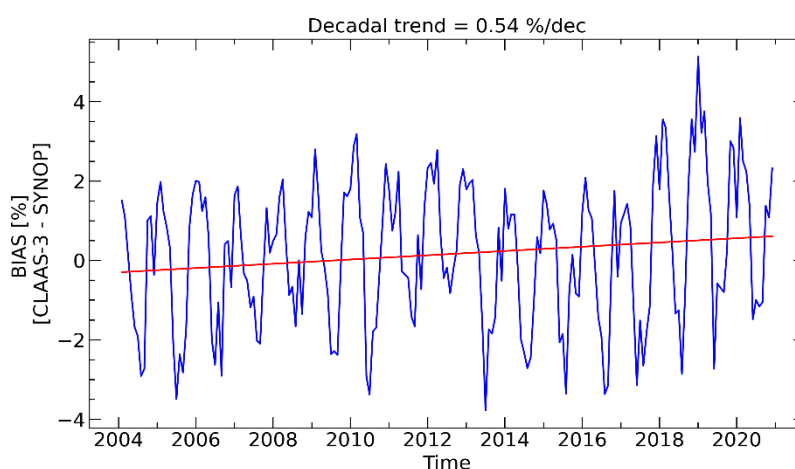


Figure 7-4: Time series of the bias between the CLAAS-3 and the SYNOP cloud fractional cover monthly mean. The red line is the linear fit.

The decadal stability gives information on the stability of the data record, for example if the data record has any unnatural trends. 17 years of CLAAS-3 data record are considered being enough to examine the decadal stability. In Figure 7-4 the temporal variation of the bias between CLAAS-3 and SYNOP CFC monthly mean and the correspondent linear fit is shown. The trend of 0.54 % per decade fulfils the optimum requirement and indicates a very stable time series, as shown in Figure 7-2.

The results of this section can be summarized as follows:

- CLAAS-3 CFC products are well within target and optimum requirements. The scores are summarized in Table 7-2.

- Monthly mean CLAAS-3 CFC values show a very small bias, 0.2 % on average, over the entire disk. The CLAAS-3 disk mean standard deviation of monthly mean CFC lies around 9 %.
- The time series of the bias is stable over the 17 years with a slight increasing trend in the last years.
- CLAAS-3 has a negative bias with respect to SYNOP for viewing zenith angle (VZA) < 60° and a positive bias for larger VZA.
- Underestimation is most significant in the western Mediterranean region and in several coastal areas. Overestimation of CLAAS-3 CFC is found in areas with a high surface albedo.

Table 7-2: Time series averages of bias and bc-rmse, and stability for CLAAS-3 CFC that were evaluated based on corresponding SYNOP monthly mean data. Remark: The anticipated error of SYNOP observations is probably of the order of 10 %, i.e., close to the Threshold requirement. Scores are color-coded: **worse than threshold**, **fulfils threshold**, **fulfils target**, and **fulfils optimal** requirements.

Parameter	Bias	bc-rmse	bias stability [/dec]
cfc [%]	0.16	8.89	0.54

7.2 Validation with CALIOP

In this section, the monthly mean cloud properties from the CLAAS-3 data record are compared against the CALIPSO-GEWEX L3 cloud product. Due to the shorter availability of the CALIPSO-GEWEX data record compared to CLAAS-3 only the overlapping period from June 2006 to December 2016 is used for the following analysis.

As described in Section 11.2, CLAAS-3 cloud parameters will be validated against two different CALIPSO flavors: top layer (uppermost-detected layer) and passive (clouds with optical thickness ≥ 0.3). Passive flavor was taken into account because it was developed with an approach to find a better agreement with the measurements of passive sensors, like SEVIRI.

The comparisons include maps, time series and zonal averages of cloud parameters. Whereas the maps present the full-disk coverage, the plots of spatially averaged values are based on data where the SEVIRI viewing zenith angle does not exceed 75°. This limitation is motivated by the high cloud amount in CLAAS-3 at the edge of the disk, as reported in Section 5.4, as well as by the consistency with validation of L2 products against CALIPSO, see Section 6.1.

Validation results of monthly mean cloud fractional cover (CFC) are shown in Section 7.2.1. The comparison of monthly means cloud top height (CTH) between CLAAS-3 and CALIPSO-GEWEX is performed in Section 7.2.2.

7.2.1 Cloud Fractional Cover (CFC)

CLAAS-3 and CALIPSO-GEWEX data records will be compared separately for total, low-level, mid-level and high-level CFC. Cloud separation is defined by use of vertical pressure levels: 680 hPa is the upper limit for low clouds and 440 hPa is the lower limit for high clouds.

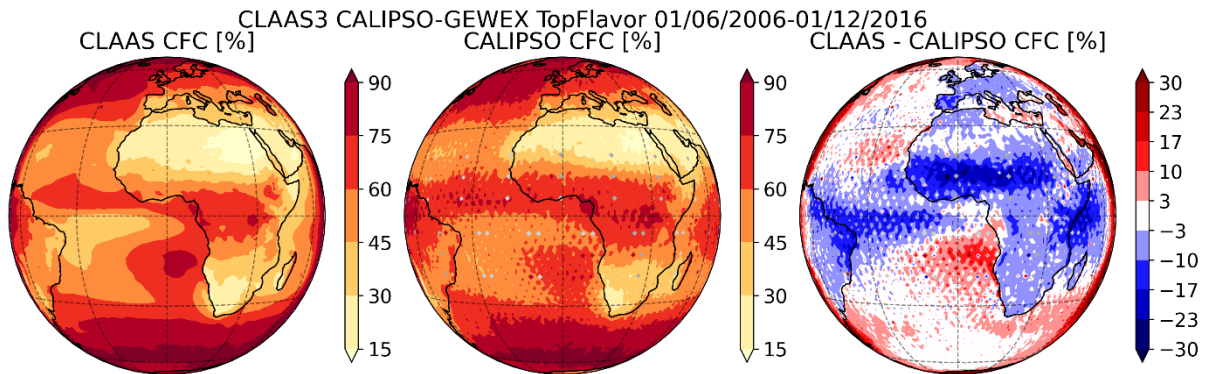


Figure 7-5: Maps of averaged cloud fractional cover based on the overlapping time period of CLAAS-3 and CALIOP L3 (2006/06 – 2016/12) from CLAAS-3 (left column), top layer flavor of CALIPSO-GEWEX cloud product (middle column) and their difference (right column).

The average monthly means of total CFC from CLAAS-3, CALIPSO-GEWEX top layer and the corresponding difference are shown in Figure 7-5. Spatial patterns for both data sets show a good agreement. The bias map, in the right panel, illustrates an underestimation of CLAAS-3 CFC for the equatorial areas, likely related to more detected thin cirrus clouds by CALIOP, Figure 7-7. An overestimation of CLAAS-3 total CFC can be found at the edge of the Meteosat disk as well as near the South-West African coast – a marine stratocumulus area. As discussed in Section 11.2, the GEWEX top layer product tends to underestimate cloud amount in regions with predominantly boundary layer cloudiness, so it makes it difficult to validate CLAAS-3 CFC against CALIPSO-GEWEX product in this particular region.

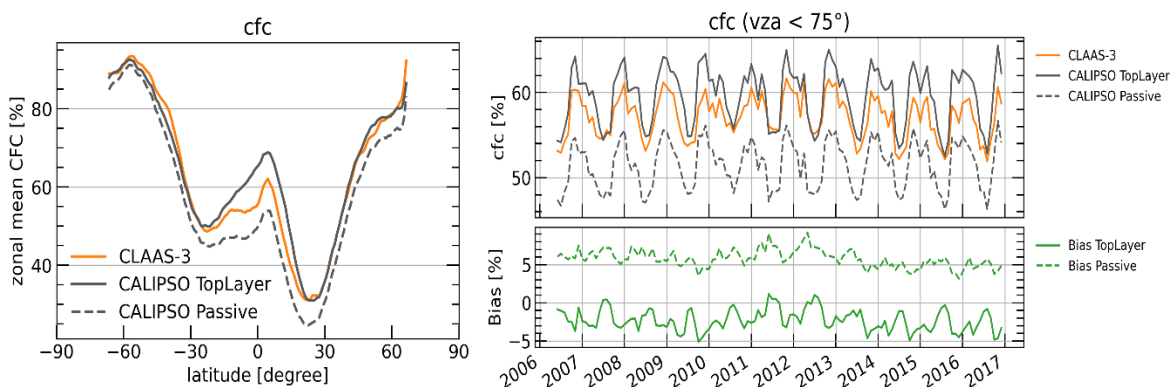


Figure 7-6: Averaged zonal mean CFC (left), and time series of CFC and bias (right) from CLAAS-3 and the CALIPSO-GEWEX top layer and passive flavor products.

Globally averaged values of total CFC from CLAAS-3 and CALIPSO-GEWEX in the upper right panel of Figure 7-6 correspond well if looking at the top layer flavor of CALIPSO data set. The difference between both time series is stable, except for a slight increase in CLAAS-3 values between 2010 and 2013, which was discussed in Section 5.1. The zonal mean plot in the left

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

panel of Figure 7-6 confirms the spatial difference patterns illustrated in the maps in Figure 7-5, when looking at the solid line: CLAAS-3 is underestimated in the tropics, the difference in the mid-latitudes is very small in average. The values with VZA higher than 75° are not presented.

Comparison with the passive flavor of CALIPSO-GEWEX shows greater difference to CLAAS-3 CFC because of the absence of optically thin clouds in the tropics. The corresponding averaged absolute bias for the passive flavor is about 2 times higher than for the top layer flavor. Considering the differences in bias values between CALIPSO flavors it can be assumed that CLAAS-3 detects a considerable part of the optically thin clouds, and the threshold for optical depth of 0.3 is likely too high to use it for the validation study. Nevertheless, passive flavor will be part of the following comparisons for consistency with the validation against CALIOP L2 data.

Figure 7-7 presents maps of temporally averaged CFC on different pressure levels for CLAAS-3 and CALIPSO-GEWEX top layer flavor and their differences. The comparison to the maps of averaged total CFC shows that the total CFC is mostly dominated by low-level and high-level CFC. The spatial patterns of all CFC subsets correspond well. The averaged low-level CFC from CLAAS-3 is higher over ocean, the mid-level CFC is higher almost over the whole domain, and high-level CFC shows a reverse behaviour.

The time series of spatially averaged CFC, right column in Figure 7-8, illustrates small differences of approximately 2 % for low- and mid-level CFC between different CALIPSO flavors. The filtering of very thin clouds in the passive flavor data set leads to higher low- and mid-level CFC. CLAAS-3 low-level CFC shows a better agreement with CALIPSO top layer flavor in winter months and with passive flavor in summer for global averaged values excluding high latitudes. CLAAS-3 mid-level CFC shows greater values than CALIPSO-GEWEX top layer flavor, so it is in a better agreement with passive flavor CFC. The corresponding averaged bias is 1.2 % for low-level CFC and 3 % for mid-level CFC. The zonal mean plot for mid-level CFC, middle left panel in Figure 7-8, clearly show overestimation of CLAAS-3 against both CALIPSO flavors outside the tropics.

The differences between top layer flavor and passive flavor for high-level CALIPSO CFC are the most prominent. The time series of CLAAS-3 high-level CFC, bottom panel in Figure 7-8, lies between the time series of top layer and passive flavors from the CALIPSO-GEWEX data set. The averaged bias is slightly higher if comparing to the passive flavor. Thus, the CLAAS-3 CFC product does not contain all optically thin high clouds detected by CALIOP but on average it includes a considerable fraction of the clouds with optical thickness less than 0.3. A detailed validation against CALIOP data with different optical thickness thresholds can be found in Section 6.1.1.

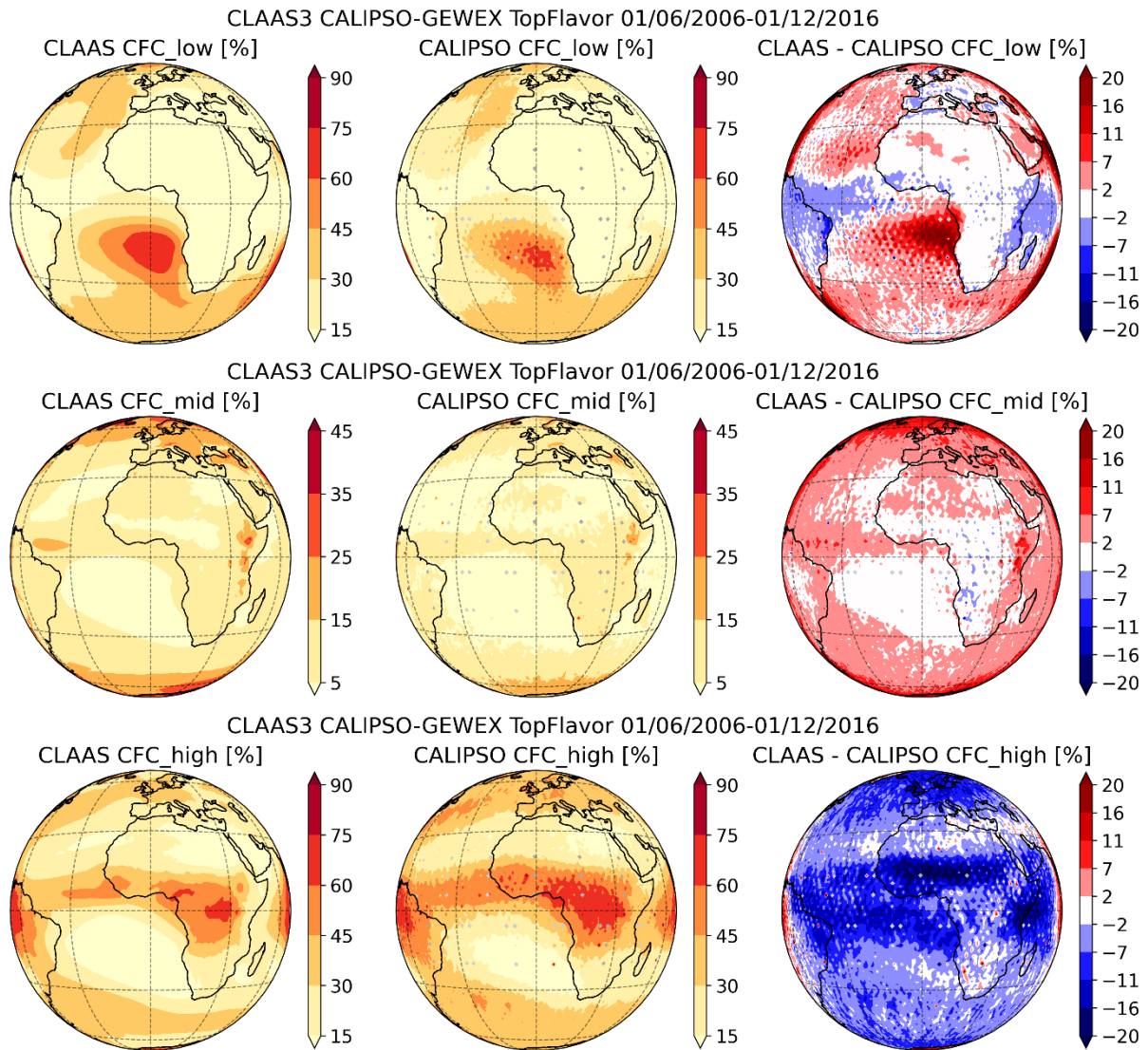


Figure 7-7: Maps of averaged fractional cover of low (top), middle (middle), and high (bottom) clouds based on the time period 2006/06 – 2016/12 from CLAAS-3 (left column), top layer flavor of CALIPSO-GEWEX cloud product (middle column) and their difference (right column).

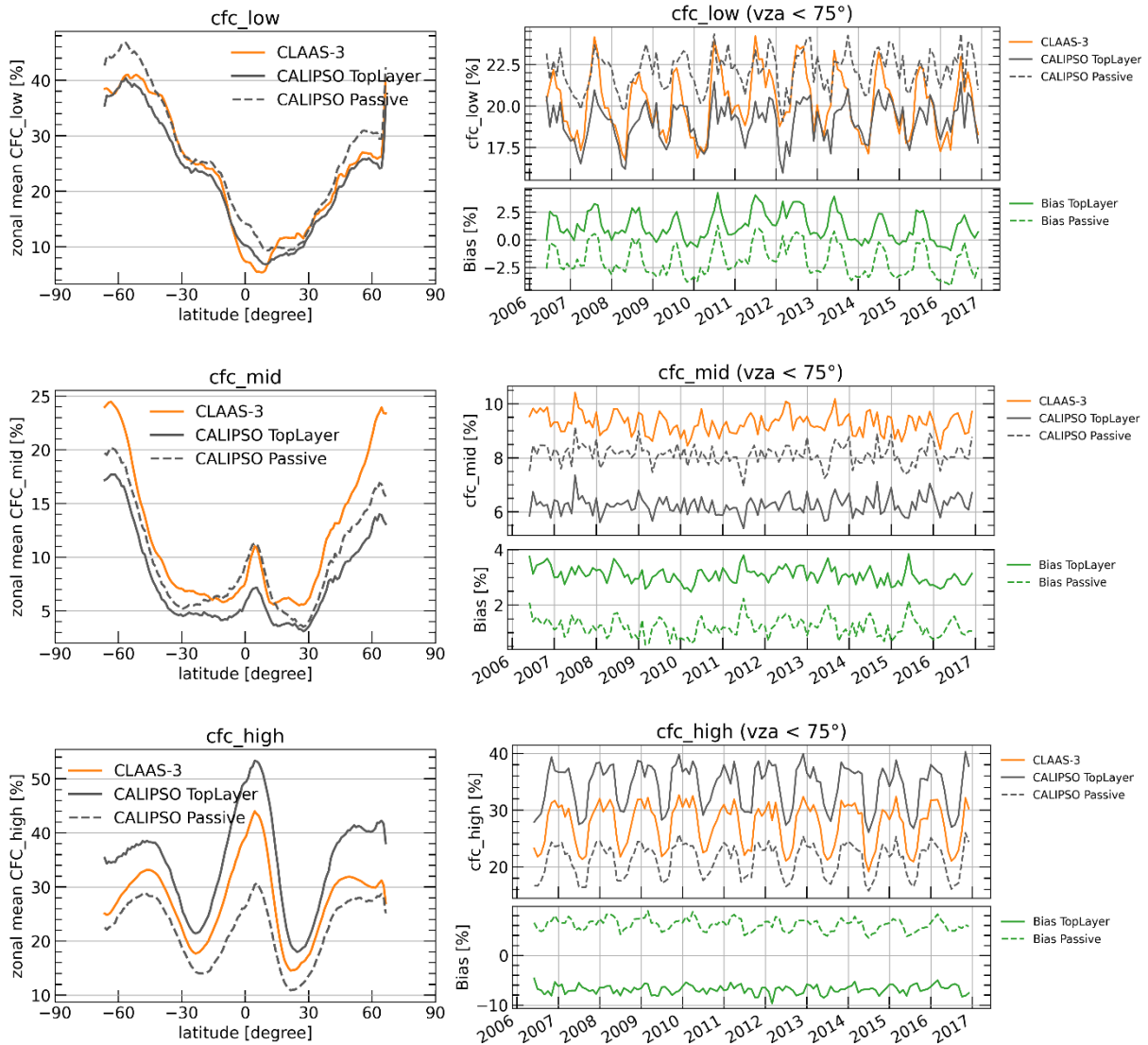


Figure 7-8: Averaged zonal mean CFC (left), and time series of CFC and bias (right) from CLAAS-3 and CALIPSO-GEWEX top layer and passive flavor for low-level clouds (upper panel), mid-level clouds (middle panel) and high clouds (bottom panel).

7.2.2 Cloud Top Height (CTH)

This section presents validation of monthly means of cloud top height (CTH) from CLAAS-3 against CALIPSO-GEWEX.

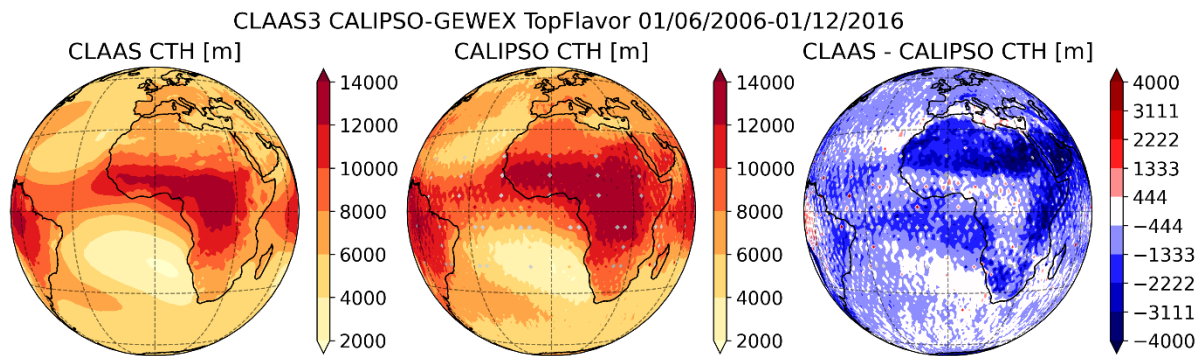


Figure 7-9: Maps of averaged cloud top height based on the time period 2006/06 – 2016/12 from CLAAS-3 (left column), top layer flavor of CALIPSO-GEWEX cloud product (middle column) and their difference (right column).

Spatial patterns of averaged CTH from CLAAS-3 and CLALIPSO-GEWEX top layer flavor are shown in Figure 7-9. The geographical distribution of the mean CTH is similar for both data sets. The CALIPSO top layer flavor shows higher cloud top height than CLAAS-3 over almost the whole domain.

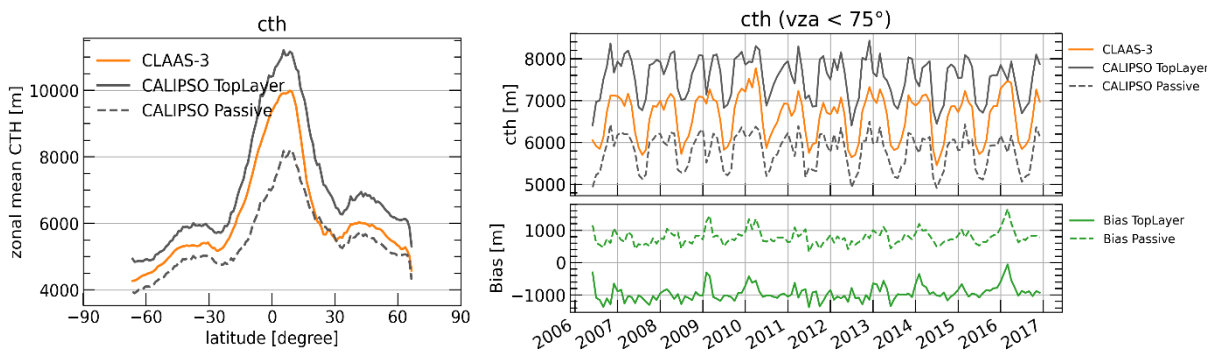


Figure 7-10: Averaged zonal mean CFC (left) and time series of CTH and bias (right) from CLAAS-3 and CALIPSO-GEWEX top layer and passive flavor.

Globally averaged time series of CTH and biases as well as zonal mean plots of CTH from CLAAS-3 and CALIPSO-GEWEX top layer and passive flavors are shown in Figure 7-10. The time series of CTH in both data records and corresponding biases are stable. The bias of passive flavor CTH is positive, so contrary to top layer there is an overestimation of the CLAAS-3 CTH. The absolute value of the averaged bias is smaller for passive flavor than for top layer CTH. Furthermore, CLAAS-3 CTH and CALIPSO passive flavor CTH in the northern mid-latitudes correspond very well.

7.2.3 Summary of results

The results of the validation of CLAAS-3 CFC and CTH against CALIPSO-GEWEX can be summarized as follows:

- CLAAS-3 CFC and CTH biases fulfil at least the threshold requirement, but bc-rmse misses the threshold requirements for CFC; biases are very stable over time and the bias stability is within the target requirement.

- Both products were compared with two different CALIPSO flavors: the CFC product is in a better agreement with the top layer flavor, CTH is in better agreement with the passive flavor.
- Low- and mid-level CFC validation results are better if comparing against CALIPSO products with filtered optical thin clouds (i.e. passive flavor).
- The most prominent spatial differences between CLAAS-3 and CALIPSO-GEWEX were found over the tropics and in the marine stratocumulus area.
- The validation scores for both CALIPSO flavors are summarized in Table 7-3. It should be noted that the bc-rms deviations are strongly impacted by the low sampling frequency of CALIOP, which explains the relative high values. Spatially smoothing the CALIPSO L3 data significantly reduces the bc_rms deviations.

Table 7-3: Time series averages of bias and bc-rmsd, and stability for total, low-, mid- and high-level CFC and CTH from CLAAS-3 data record that were evaluated based on corresponding CALIPSO-GEWEX L3 data. Scores are color-coded: **worse than threshold**, **fulfils threshold**, **fulfils target**, and **fulfils optimal** requirements.

Parameter	CALIPSO top layer flavor			CALIPSO passive flavor		
	Bias	bc-rmsd	bias stability [°/dec]	Bias	bc-rmsd	bias stability [°/dec]
cfc [%]	-2.29	22.52	-0.79	5.79	21.02	-1.52
cfc_low [%]	1.22	16.71	-0.53	-1.85	17.13	-0.51
cfc_middle [%]	3.08	11.27	-0.27	1.21	12.59	-0.05
cfc_high [%]	-6.84	24.66	0.04	6.19	20.76	-0.93
cth [m]	-951	3319	178	780	3043	73

7.3 Validation with MAC-LWP

The MAC-LWP dataset provides monthly mean all-sky LWP in $1^\circ \times 1^\circ$ grid boxes over ocean that is based on all available data for a specific month (see Section 11.5). In addition, for each month and each grid box over the 1988-2016 period, monthly climatologies of diurnal cycle amplitudes and phases are provided. These are used to obtain the monthly mean all-sky LWP in 1-hour time steps, to compare it with corresponding monthly mean diurnal cycle CLAAS-3 LWP. The relevant calculations are based on the equation:

$$LWP(Y, t) = \overline{LWP(Y)} + A_1 \cos(\omega(t - T_1)) + A_2 \cos(2\omega(t - T_2))$$

where $\overline{LWP(Y)}$ is the uncorrected monthly mean LWP for year Y , t is the 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 Elsaesser et al. 2017).

As described in Section 11.5, the marine stratocumulus region in the southeastern Atlantic off the Namibian coast was selected for this validation (Figure 7-11). Monthly mean diurnal LWP values were computed as area-weighted spatial averages from this region and span the 02/2004-12/2016 period, when both CLAAS-3 and MAC-LWP data sets are available. It should be noted that, since MAC-LWP provides all-sky LWP data, all-sky monthly mean diurnal LWP was also used from CLAAS-3. Since this parameter is not directly available in the CLAAS-3 monthly-mean diurnal cycle files, it was calculated by multiplying the monthly mean diurnal in-cloud LWP (which is based on the 3.9 μm retrievals) with the corresponding CFC and CPH (liquid cloud fraction). Additionally, only time slots when the region was fully covered by both data sets were considered. Monthly CPH was found to be always above 92%, verifying the assumption that almost exclusively liquid clouds are present in this region.

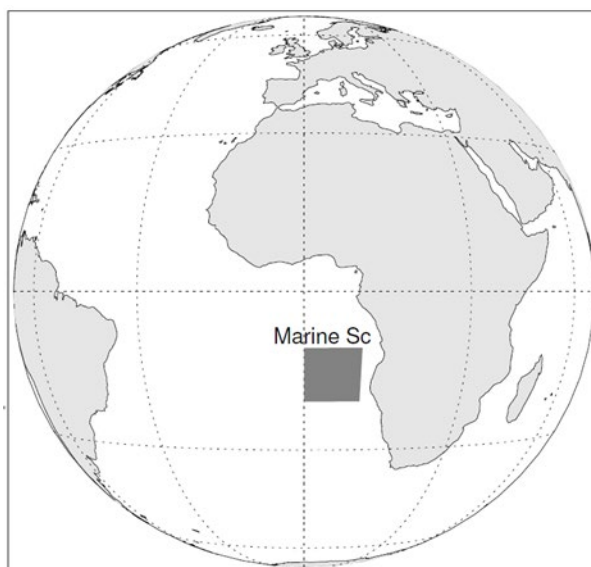


Figure 7-11: The marine stratocumulus region off the Namibian coast (0-10°E, 10-20°S), selected for the validation of CLAAS-3 all-sky LWP monthly mean diurnal cycle against corresponding MAC-LWP data.

Figure 7-12 shows the average time series plots of the monthly mean all-sky LWP between 7:00 UTC and 16:00 UTC from CLAAS-3 and MAC-LWP, along with their bias and bc-rmsd. The agreement between the two data sets is very good, with similar seasonal characteristics. The bias fluctuates in the -20 to 20 g m^{-2} zone with an average value of 0.26 g m^{-2} . Bias maxima and minima occur due to higher CLAAS-3 values in August to November and lower in January to April. A solid explanation for this seasonality in the bias is hard to give. The bc-rmsd fluctuates much less around an average value of 11.37 g m^{-2} .

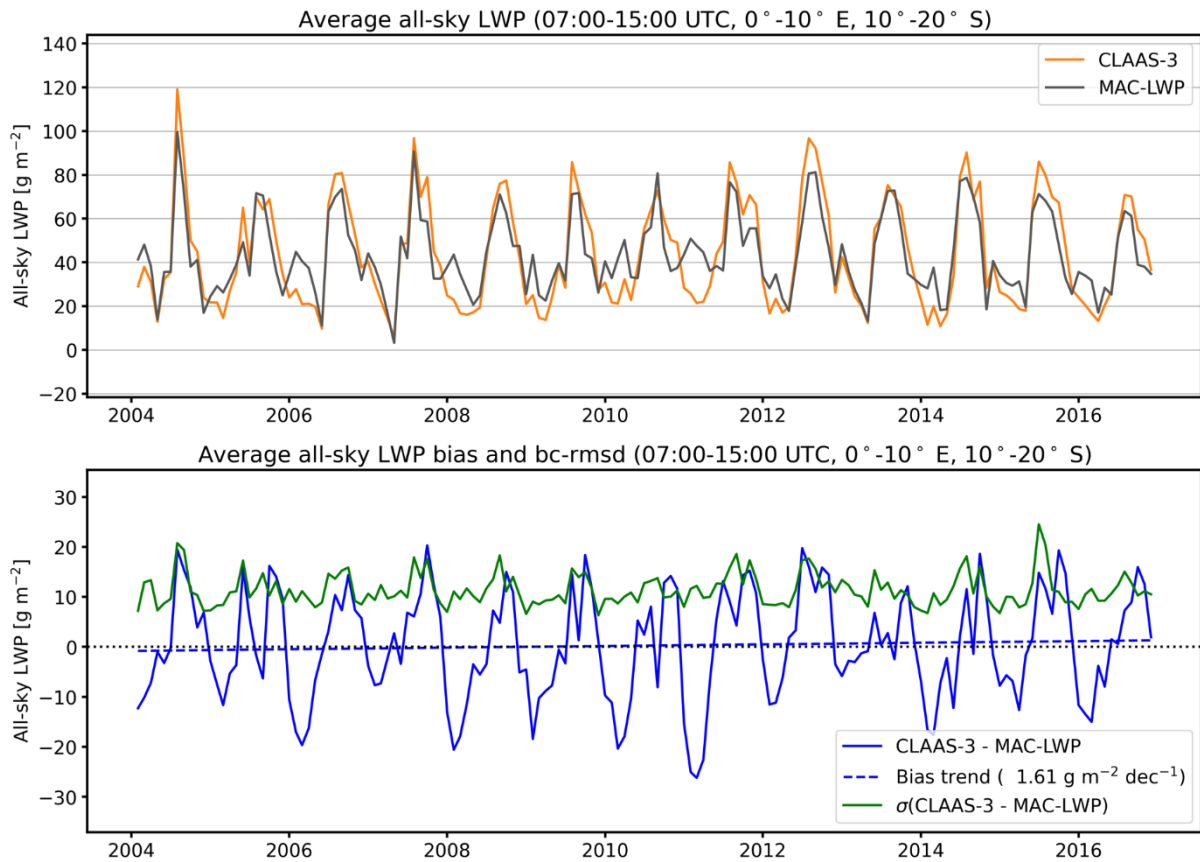


Figure 7-12: Time series of the spatially averaged all-sky LWP at the marine Sc region (0-10°S, 10-20°E) from CLAAS-3 and MAC-LWP data (top plot). The values are averages of the 1-hour time slots between 7:00 UTC and 16:00 UTC. The corresponding bias and bc-rmsd are shown in the bottom plot, along with the bias trend during the period examined.

The monthly mean diurnal variation of the all-sky LWP from CLAAS-3 and MAC-LWP is shown in Figure 7-13. In order to ensure the representativeness of results in the region, spatial averages were computed only when all grid cells had valid LWP values. Due to this requirement, nighttime hours (obviously absent in CLAAS-3 LWP retrievals) and hours early in the morning and late in the afternoon are excluded from this analysis. In the 9-hour part of the diurnal cycle depicted in Figure 7-13, both data records show a reduction in all-sky LWP throughout the day, except for the last hour in CLAAS-3. Another interesting feature is that the bias is below 10 g m⁻² in all but the first time slot, fulfilling the target requirement. The bc-rmsd is close to the 10 g m⁻² optimal requirement most of the times.

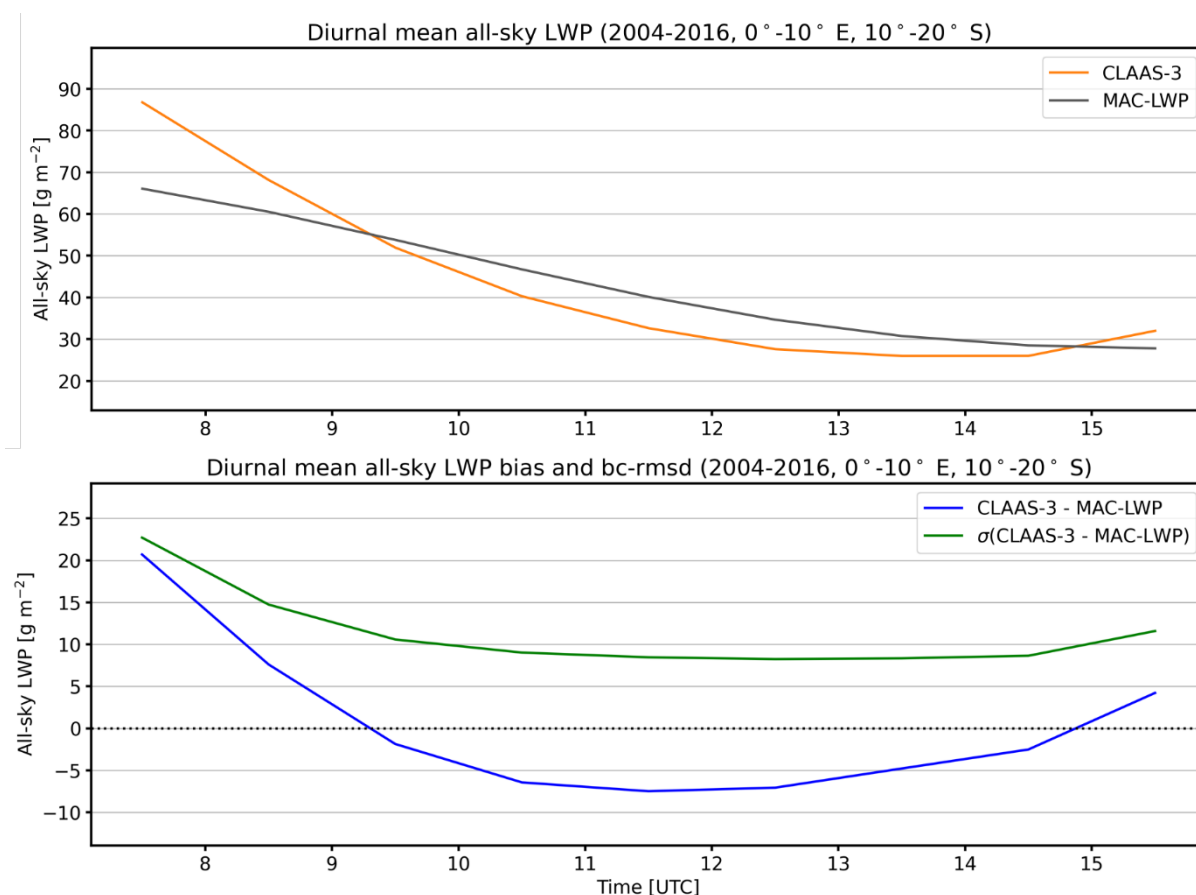


Figure 7-13: Monthly mean diurnal average of all-sky LWP from CLAAS-3 and MAC-LWP (top plot), spatially averaged at the marine Sc region (0-10°S, 10-20°E), along with their bias and bc-rmsd (bottom plot).

Table 7-4 summarizes the overall validation results for CLAAS-3 all-sky LWP in terms of the bias and bc-rmsd with respect to the corresponding MAC-LWP product. Results correspond to the 7:00 UTC – 16:00 UTC average values, calculated from the 2004-2016 period, when both data products are available. It is found that both the bias and the bc-rmsd fulfill the target requirements, with values lying close to the optimal requirements.

Table 7-4: Overall requirement compliance of the CLAAS-3 all-sky LWP product with respect to the bias, bc-rmsd, and stability of the bias, calculated over the marine stratocumulus region (10°-20° S, 0°-10°E) for validation with MAC-LWP data. Scores are color-coded: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.

Parameter	bias [g m ⁻²]	bc-rmsd [g m ⁻²]	stability of bias [g m ⁻² /decade]
All-sky LWP	0.26	11.37	1.61

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

7.4 Comparison with MODIS

The MODIS L3 Collection 061 monthly mean data set (data type names: MOD08_M3 and MYD08_M3 for Terra and Aqua, respectively), available at $1^\circ \times 1^\circ$ spatial resolution, has been used as the main reference for the cloud properties evaluation. Five aspects are important to mention here:

1) CLAAS-3 monthly mean properties are an average of all daytime SEVIRI time slots ($SZA < 75$ degrees). On the other hand, there are two MODIS instruments with equatorial overpass times at 10:30 and 13:30 local time. Monthly means from these two instruments were averaged in order to best mimic the CLAAS-3 averaging. While some differences will originate in this different temporal sampling, this is not expected to cause major issues, as was shown before [RD 5].

2) CLAAS-3 optical properties (CWP, COT, CRE) are retrieved based on two channel pairs: $0.6 \mu\text{m} - 1.6 \mu\text{m}$ and $0.6 \mu\text{m} - 3.9 \mu\text{m}$, with the latter being a new feature compared to CLAAS-2.1. These retrievals are compared separately, with MODIS retrievals from corresponding channels at $1.6 \mu\text{m}$ and $3.7 \mu\text{m}$.

3) As the MODIS products, the CLAAS-3 data sets contain cloudy sky monthly averages of liquid/ice water paths and corresponding optical thickness and effective radius. These averages have been compiled in the same way (i.e., directly from the L2 data), which is essential for the products to be compared. The MODIS C61 cloud products contain partly cloudy retrievals, denoted as “PCL”, separately from the “standard” overcast pixels. The former retrievals regard cloud edges or broken clouds, and they are considered of lower quality compared to the overcast pixels. However, their exclusion from the analysis would probably lead to MODIS averages biased toward optically thicker clouds (Platnick et al. 2017), and thus the PCL pixels were included in the comparisons. For cloud water path its all-sky version was evaluated, calculated for MODIS by multiplying the in-cloud water path with corresponding cloud fraction and phase. In contrast, for cloud optical thickness and effective radius, in-cloud averages were compared.. Cloud thermodynamic phase was evaluated in terms of the fraction of liquid clouds relative to the total cloud fraction.

4) CDNC is compared against the corresponding Bennartz and Rausch climatology. While still a L3 product based on MODIS, it is produced separately from the other MODIS products. It covers the period 2004-2017, based on L2 products from MODIS Aqua only over ocean, and using the $0.6 \mu\text{m} - 3.7 \mu\text{m}$ spectral channel pair (Bennartz and Rausch, 2017).

5) Cloud optical properties retrievals rely on measurements of reflected sunlight. This implies that retrievals have to be discontinued if the solar zenith angle reaches a certain threshold. This threshold is set to 75 degrees for CLAAS-3 daytime L3 products, but has a different value (81.36 degrees) for MODIS. Similarly, CLAAS-3 L3 data sets are computed only for SEVIRI viewing zenith angles up to 75 degrees. Therefore, to enable a meaningful comparison of time series, spatial aggregation was not performed over the full SEVIRI disk, but over a subset of the disk, in which valid SEVIRI retrievals are available for all seasons and during the whole record. This restriction determined the choice of an aggregation area bounded by 45 degrees West-East and South-North.

7.4.1 Cloud Fractional Cover (CFC)

The comparison of monthly L3 cloud fractional coverage between CLAAS-3 and MODIS is performed separately for CFC, CFC_Day and CFC_Night variables. Corresponding MODIS variables are Cloud_Fraction_Mean_Mean, Cloud_Fraction_Day_Mean_Mean and Cloud_Fraction_Night_Mean_Mean, derived from L2 Cloud Mask data, by counting the lowest two clear sky confidence levels (cloudy and probably cloudy) and dividing by the total count.

Maps of time series averages from CLAAS-3, MODIS and their differences are shown in Figure 7-14. A requirement for these averages to be computed from at least 90% of the time series was always met by both data records. Spatial patterns are overall similar, with CLAAS-3 being lower than MODIS mainly around the tropics, and higher in the disk edge. The latter differences should be attributed to the challenging viewing conditions of SEVIRI at the edge of the disk, which lead to an overestimation in cloud fraction. On the contrary, closer to the centre of the SEVIRI disk MODIS samples clouds from on average higher viewing angles than SEVIRI, which may contribute to higher cloud fraction retrievals.

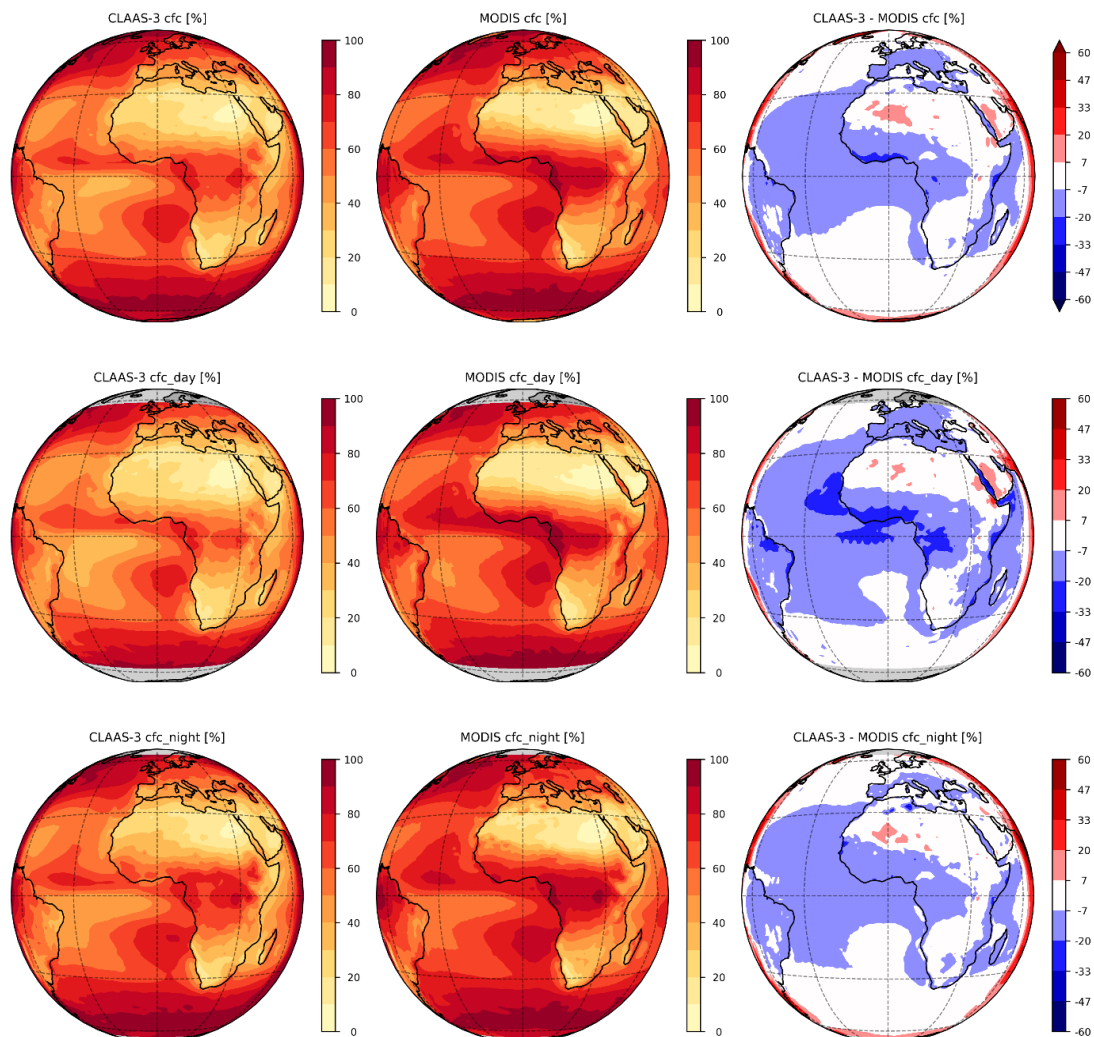


Figure 7-14: Maps of average CFC (top row), CFC_Day (middle row) and CFC_Night (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12) from CLAAS-3 (left column), MODIS (middle column) and their difference (right column).

Time series of CFC, CFC_Day and CFC_Night from CLAAS-3 and MODIS monthly L3 data, along with their biases and bc-rmsd are shown in Figure 7-15. Average values are calculated for an area from 45° S to 45° N, and from 45° W to 45° E, to ensure that polar night gaps are avoided and the same area is always considered. While the MODIS time series is overall stable, CLAAS-3 shows an increase in CFC in 2010-2013 and after 2017, which should be attributed mainly to CFC_Night retrievals. Examination of average maps from 2011-2012 (not shown) indicates that this increase is spatially uniform. A small decrease in 2019-2020, compared to previous years, is also apparent in both data sets. Spatial coverage remained complete throughout the time series. Instead, these fluctuations may be related to instabilities in the calibration of the 3.9 μm channel relative to the 10.8 μm channel (see Appendix A). Biases are stable in all three cases (blue dotted lines in the bias plots of Figure 7-15), with changes of less than 1% per decade, except for the CFC_Night bias, where the change is 1.31% dec⁻¹. This, however, should be attributed to the irregularities mentioned before.

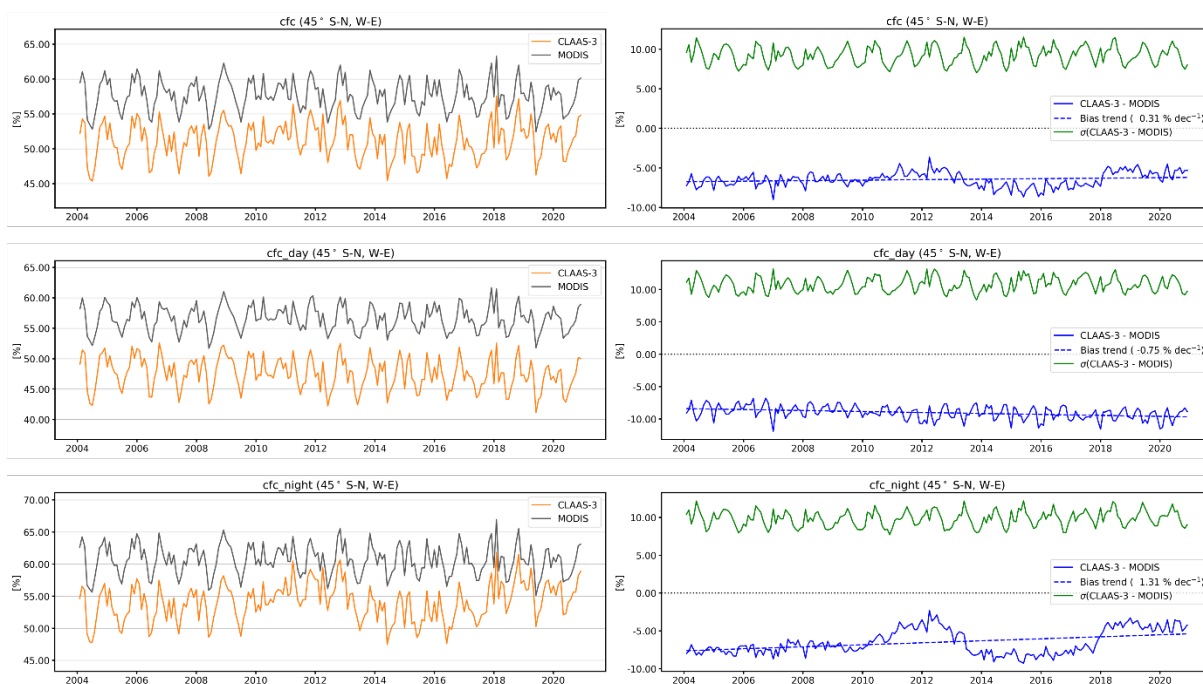


Figure 7-15: Time series of CFC (top row), CFC_Day (middle row) and CFC_Night (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (% per decade) based on linear regression.

7.4.2 Cloud Top Level

Both cloud top height (CTH) and cloud top pressure (CTP) were compared with corresponding MODIS L3 monthly data sets (Cloud_Top_Height_Mean_Mean and Cloud_Top_Pressure_Mean_Mean, respectively). Results show consistently higher cloud top

heights in CLAAS-3, which naturally lead to lower values for cloud top pressure, compared to MODIS.

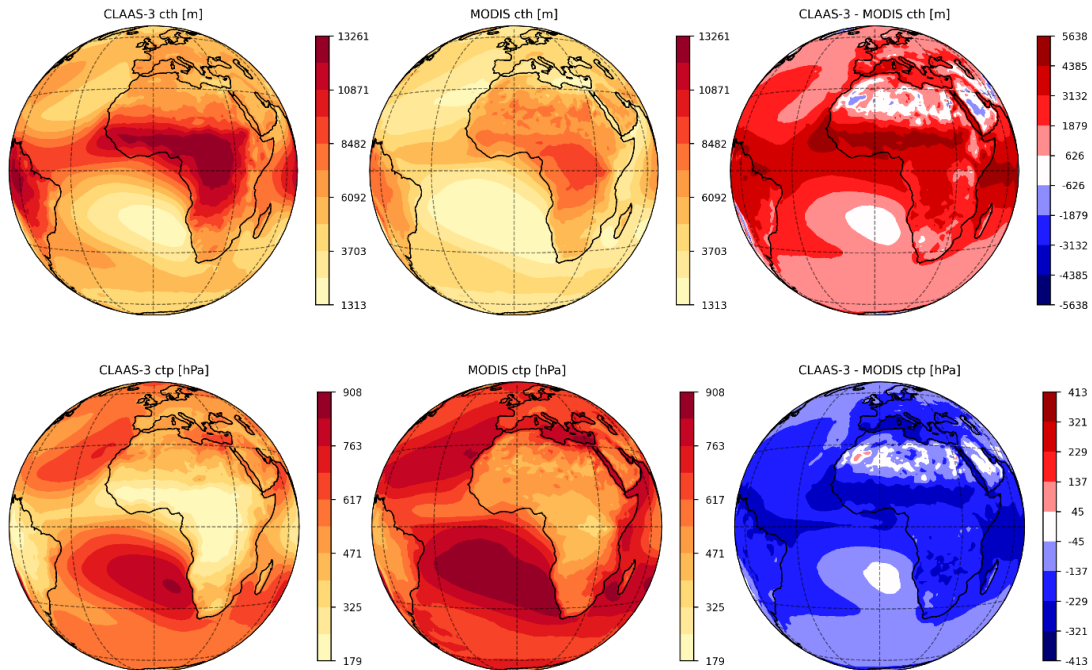


Figure 7-16: Maps of average CTH (top row) and CTP (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12) from CLAAS-3 (left column), MODIS (middle column) and their difference (right column).

The time series of both variables in both data records are very stable, leading to also stable biases, where only a seasonal fluctuation appears (Figure 7-17). The large biases (typically over 2000 m) should be attributed to the different retrieval approach of CLAAS-3 compared to MODIS (see [RD 3]). A combined inspection of Figure 7-10 and Figure 7-17 shows that CLAAS-3 CTH lies between CALIPSO and MODIS, indicating a good product quality. The improved performance of the CLAAS-3 neural network approach compared to MODIS was also shown in Håkansson et al. (2018) for different flavors of the trained network.

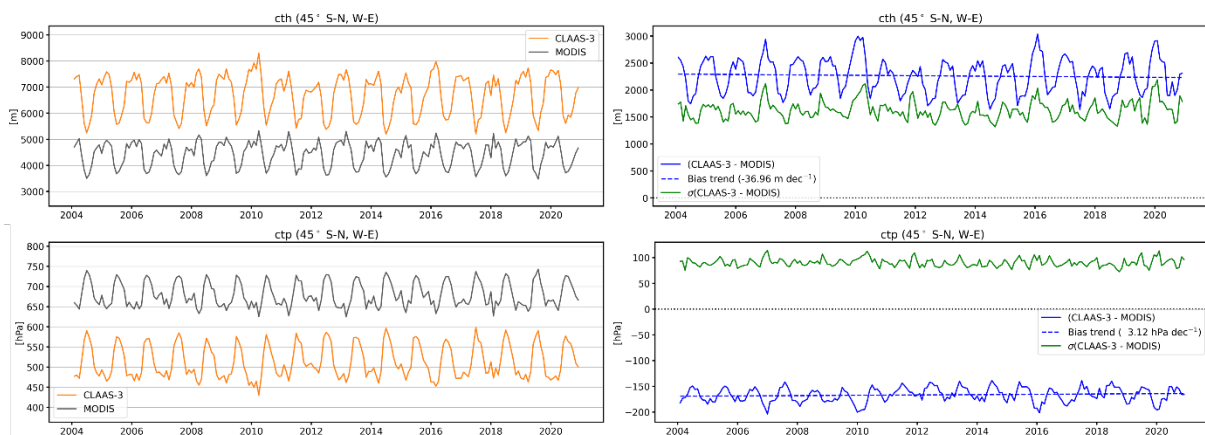


Figure 7-17: Time series of CTH (top row) and CTP (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are

averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (m and hPa per decade) based on linear regression.

7.4.3 Cloud Thermodynamic Phase (CPH)

Data records of cloud phase (CPH) and day-only cloud phase (CPH_Day) were evaluated separately against corresponding MODIS data (Cloud_Phase_Infrared_Histogram_Counts and Cloud_Phase_Infrared_Day_Histogram_Counts). In these MODIS products, retrievals are aggregated in four bins: clear, liquid, ice, and mixed or undetermined. In order to compare with CLAAS-3 CPH (which is the fraction of liquid clouds to the total clouds), MODIS cloud phase was estimated by dividing the number in bin “liquid” by the sum of numbers in bins “liquid”, “ice” and “mixed/undetermined”.

Comparisons were also performed separately for CLAAS-3 data records CPH/CPH_Day and CPH_16/CPH_16_Day. The two pairs correspond to final adjustments in the “initial” CPH retrieval, after applying the 3.9 μm (first pair) or the 1.6 μm (second pair) retrieval of COT and CRE. Here only results from the CPH/CPH_Day retrievals are shown, since the corresponding results from the 1.6 μm retrievals have very similar patterns, although the 1.6 μm liquid fraction is overall somewhat smaller. Averages of biases, bc-rmsd and bias stabilities are reported for all cases in Table 7-5.

Time series averages of CPH and CPH_Day from CLAAS-3, MODIS and their differences are shown in Figure 7-18. Spatial patterns are very similar, with CLAAS-3 acquiring values lower than MODIS near the tropics and higher over the Middle East, Europe and in higher latitudes. Differences typically range between 5 % and 15 %, exceeding 25 % in several cases.

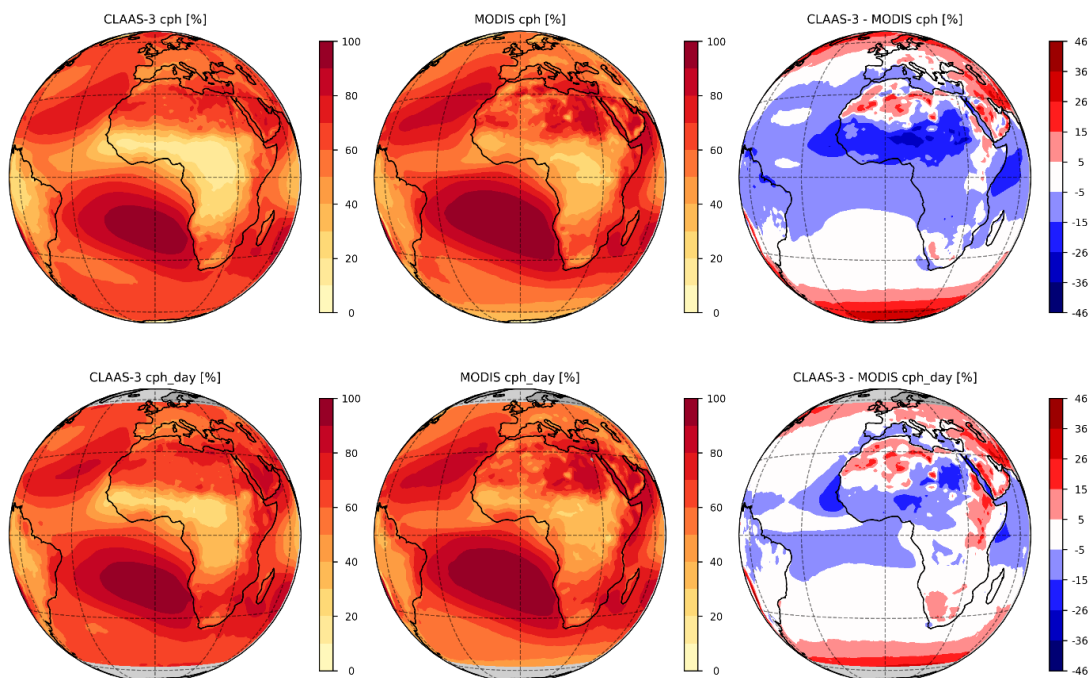


Figure 7-18: Maps of average cloud phase (CPH, top row) and day-only cloud phase (CPH_Day, bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12) from CLAAS-3 (left column), MODIS (middle column) and their difference (right column).

Time series of CLAAS-3 and MODIS CPH and CPH_Day (spatial averages in the 45° S-N and W-E region) and their biases and bc-rmsd are shown in Figure 7-19. In both cases biases get closer to zero after 2012. A closer examination of the time series plots shows that this should be attributed to a slight decrease in corresponding MODIS values. It should be noted, however, that this feature causes an increase in the slopes of the relevant bias regressions, which leads to a decadal change in the bias of almost 2.5% and 3% for CPH and CPH_Day, respectively.

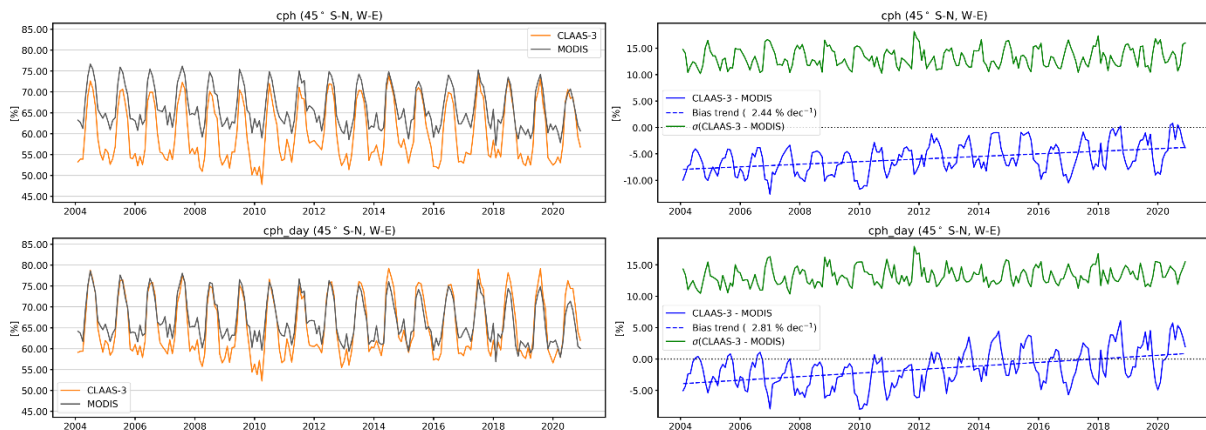
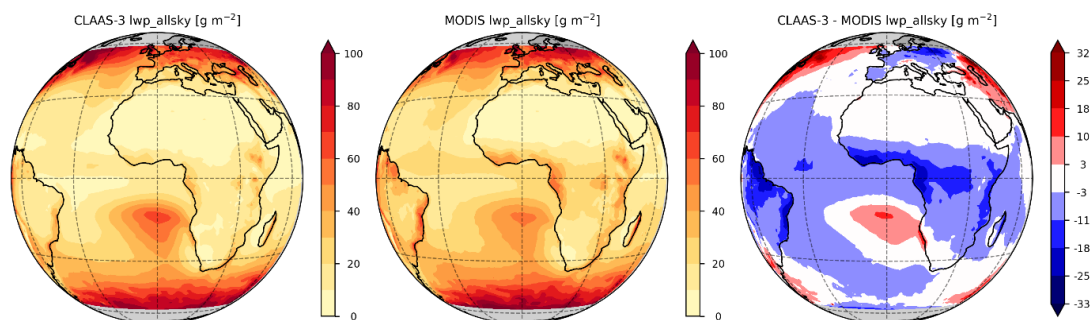


Figure 7-19: Time series of CPH (top row) and CPH_Day (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (% per decade) based on linear regression.

7.4.4 Liquid Water Path (LWP)

Average maps of the all-sky LWP are shown in Figure 7-20, separately for the retrievals from the 3.9 μm and the 1.6 μm channels. While the spatial features are similar, with higher values appearing mainly in higher latitudes and secondarily in the stratocumulus region of the southeastern Atlantic, the CLAAS-3–MODIS difference patterns vary between the two retrievals. As shown in Section 7.4.4.1, spatial features are dominated by corresponding features in liquid COT, while the spatially averaged differences originate in corresponding differences in liquid CRE.



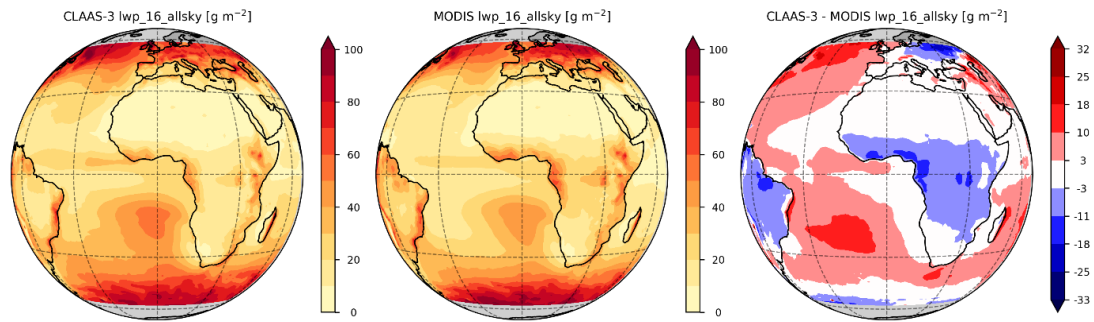


Figure 7-20: Maps of average all-sky Liquid Water Path (LWP) retrieved based on the 3.9 μm channel (top row) and on the 1.6 μm channel (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left column), MODIS (middle column) and their difference (right column).

The time series of the all-sky LWP in both channel pair retrievals show good agreement with MODIS, in terms of both seasonality patterns and absolute values. In terms of bias stability, both comparisons produce very stable results, although a small dip appears between 2010 and 2013 in the bias of the 1.6 μm retrieval.

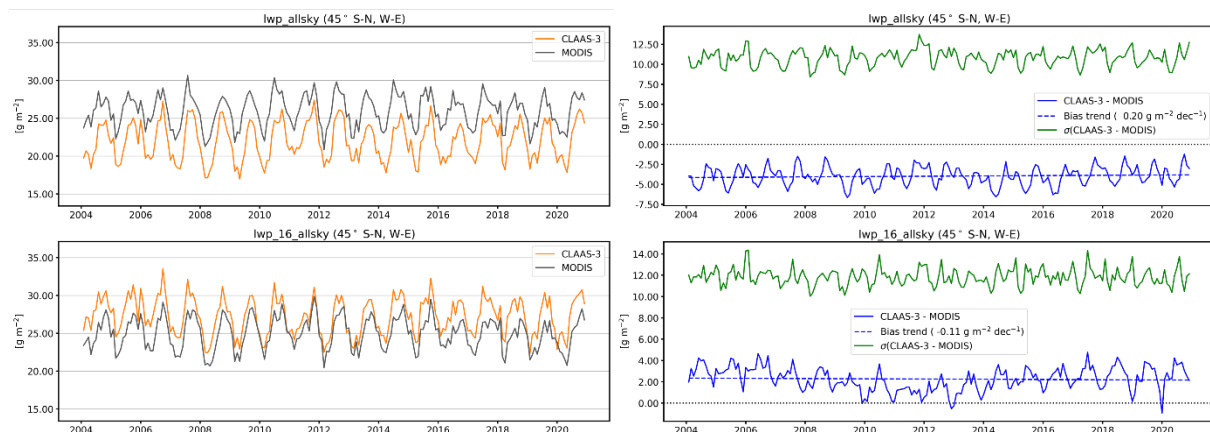


Figure 7-21: Time series of all-sky LWP retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (g m^{-2} per decade) based on linear regression.

7.4.4.1 Liquid COT and CRE

As mentioned before, spatial patterns in the all-sky LWP are driven mainly by liquid COT. This is verified in Figure 7-22, where the similarity is obvious. High liquid COT values at higher latitudes should be partly attributed to illumination conditions, with higher solar zenith angles leading to larger COT retrieved. This seems more pronounced in MODIS (see Figure 6-23). There is no apparent difference between the 3.9 μm and the 1.6 μm liquid COT retrievals and their MODIS counterparts. For this reason the 1.6 μm retrievals are not shown here. In both cases CLAAS-3 retrieves slightly higher values over land near nadir and in some areas of Europe and Asia, which are probably affected by snow on the underlying surface. In most other cases, however, and especially over land, MODIS retrieves higher values.

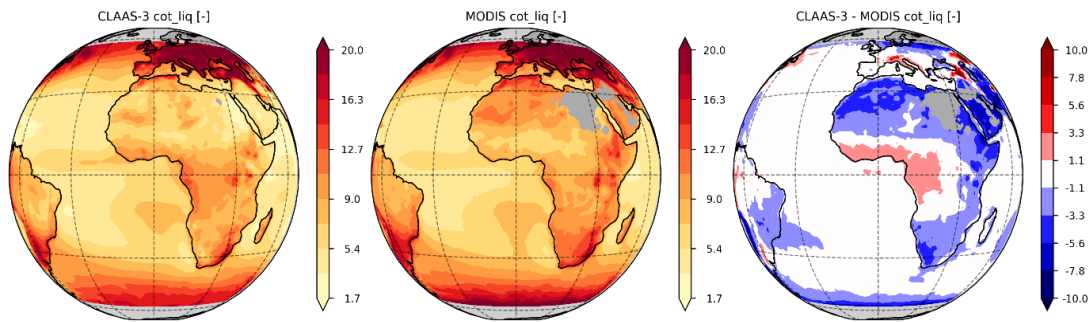


Figure 7-22: Maps of average liquid Cloud Optical Thickness (COT) retrieved based on the 3.9 μm channel, and on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left column), MODIS (middle column) and their difference (right column).

On the other hand, there are noticeable differences between the 3.9 μm and the 1.6 μm liquid CRE retrievals, shown in Figure 7-23. While spatial features are similar, the 1.6 μm retrieval is on average larger, with differences being more pronounced in CLAAS-3 compared to MODIS. In fact, CLAAS-3 is on average higher than MODIS in the 1.6 μm retrieval and lower in the 3.9 μm (Figure 7-25). This leads to the different patterns of CLAAS-3 – MODIS differences (right column in Figure 7-23): while larger CLAAS-3 values prevail in the 1.6 μm , there is a latitudinal gradient in the 3.9 μm case.

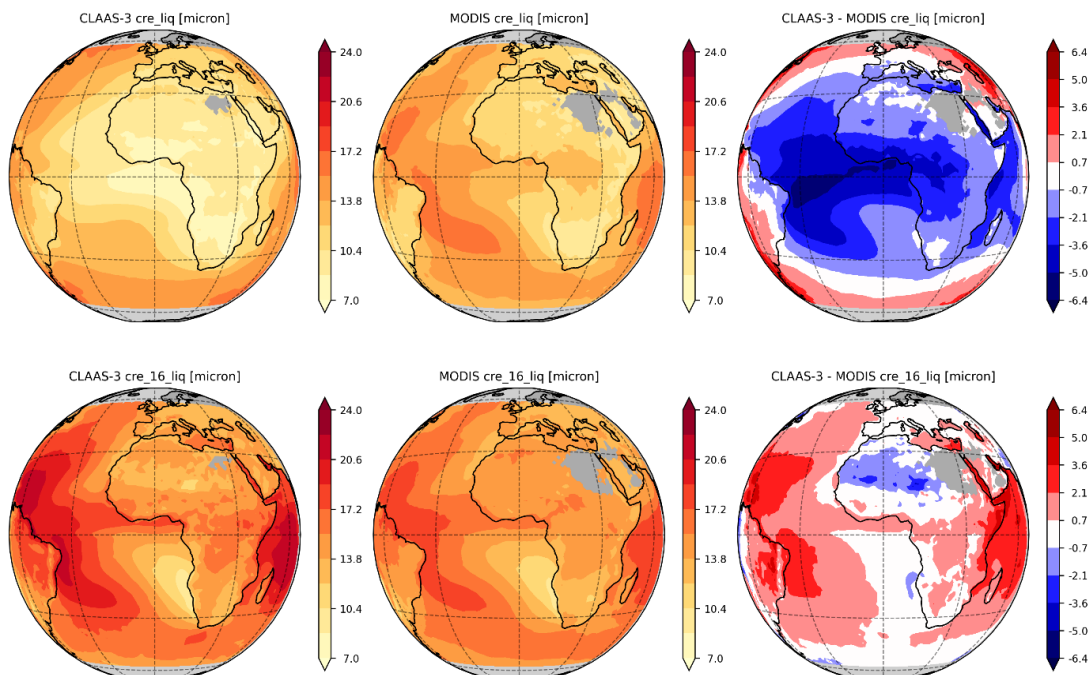


Figure 7-23: Maps of average liquid Cloud Effective Radius (CRE) retrieved based on the 3.9 μm channel (top row) and on the 1.6 μm channel (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left column), MODIS (middle column) and their difference (right column).

Time series of COT are very stable in both CLAAS-3 and MODIS, leading to stable biases throughout the CLAAS-3 period (Figure 7-24). On the other hand, there are some minor irregularities in the CLAAS-3 3.9 μm liquid CRE, which seem to be related to the 3.9 μm

channel relative calibration issue also affecting CFC (see Section 7.4.1 and Appendix A). Both data records show, for unclear reasons, irregularities in the 1.6 μm liquid CRE, which also leads to a noticeable decadal change in the bias. It can be seen in Figure 7-24 that the CLAAS-3 3.9 μm (in-cloud) COT is smaller than the 1.6 μm COT. The opposite is true for ice clouds (see Section 7.4.5.1). It was verified that this is mainly due to differences in CPH, namely more liquid clouds in the 3.9 μm flavor leading to a smaller average liquid COT, while on a pixel-to-pixel basis COT is very similar between the two retrievals (see also Section 6.4.2). Average values of liquid COT and CRE biases, bc-rmsd and bias stability with respect to MODIS are reported in Table 7-5.

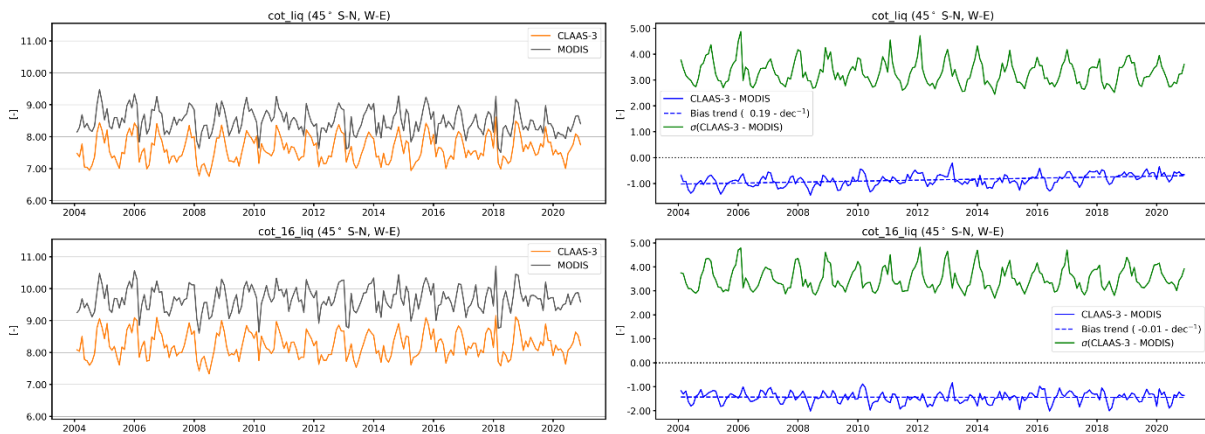


Figure 7-24: Time series of liquid COT retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (per decade) based on linear regression.

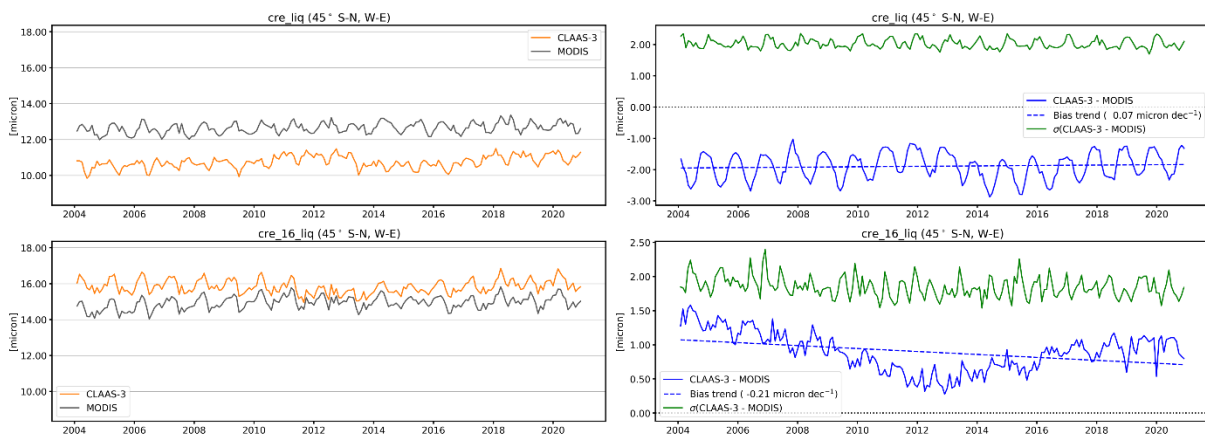


Figure 7-25: Time series of liquid CRE retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (μm per decade) based on linear regression.

7.4.4.2 Liquid CDNC

Spatial distributions of CLAAS-3 and MODIS liquid CDNC, both based on the 3.9 μm retrievals and the common period 2004/02-2015/12 are shown in Figure 7-26. Two comparison options were examined: a) averages in all grid cells where the time series is at least 50% complete, separately from CLAAS-3 and MODIS (top plots in Figure 7-26); b) averages based only on cases where both CLAAS-3 and MODIS are available (bottom plots in Figure 7-26). In the latter case, CLAAS-3 retrievals over land are excluded, since these are not available from MODIS. In both cases, however, the pattern of differences is similar, with higher CLAAS-3 values near the disk centre and lower close to the edges, and over the Mediterranean Sea.

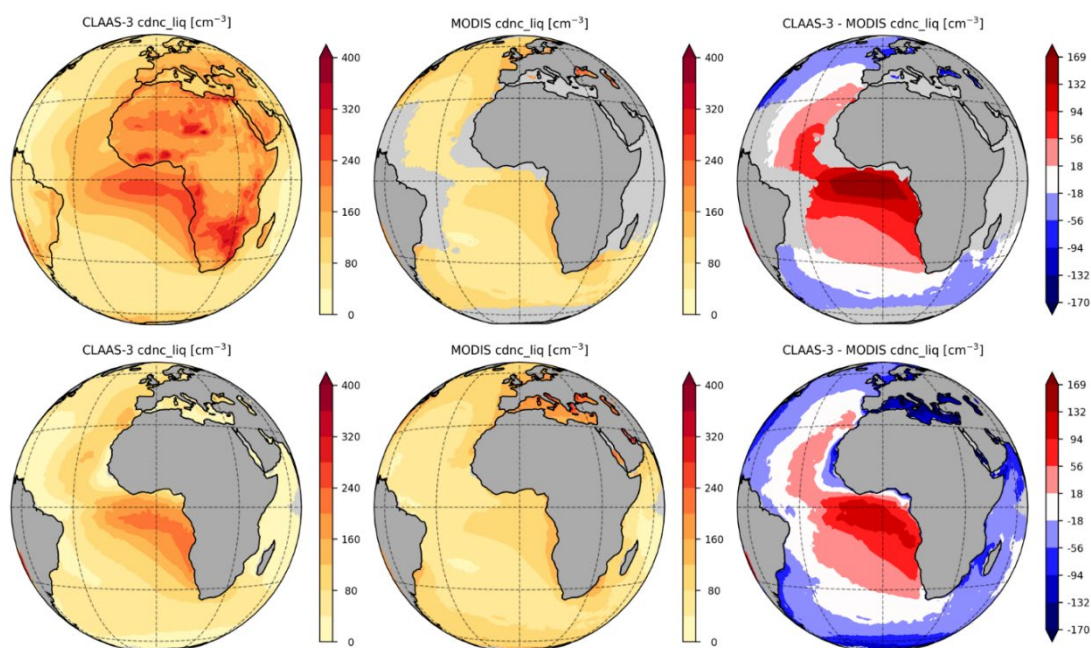


Figure 7-26: Maps of average liquid Cloud Droplet Number Concentration (CDNC) based on the period 2004/02 – 2015/12, from CLAAS-3 (left column), MODIS (middle column) and their difference (right column). In the top row, all available data are used, with the requirement that the time series is at least 50% complete on a grid cell basis. In the bottom row, only grid cells where both CLAAS-3 and MODIS are available were used in the averaging.

On a monthly basis, CLAAS-3 CDNC is well correlated with MODIS but constantly higher by about ~40% on average, which leads to an average bias of $\sim 14 \text{ cm}^{-3}$ (Figure 7-27). This difference should be attributed to differences mainly in CRE and secondarily in COT: CLAAS-3 liquid CRE and COT (based on the 3.9 μm retrievals) are approximately 15% and 10% lower than MODIS, respectively. Based on the dependency of the CDNC computation on these parameters [RD 4], these differences would lead to 50% higher and 6% lower CDNC, respectively, leading to a difference close to the one reported here.

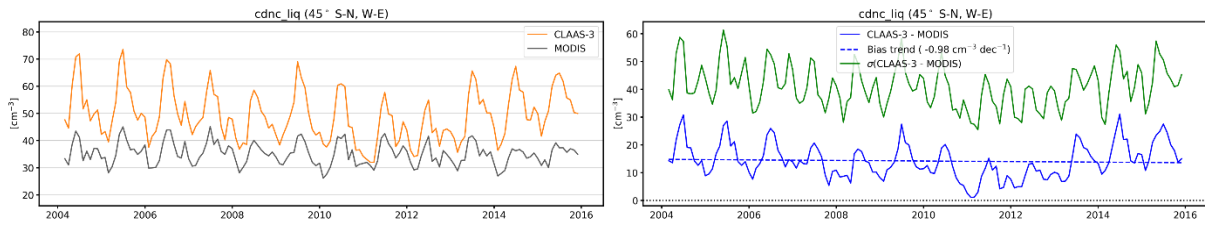


Figure 7-27: Time series of liquid Cloud Droplet Number Concentration (CDNC), retrieved based on the 3.9 μm channel, from monthly L3 CLAAS-3 and MODIS data (left plot) and corresponding bias and bc-rmsd (right plot). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E, and over sea only, due to lack of MODIS data over land. The blue dotted line in the bias plot shows the change in bias (cm^{-3} per decade) based on linear regression.

7.4.5 Ice Water Path (IWP)

Figure 7-28 shows average maps of the all-sky IWP separately for the retrievals from the 3.9 μm and the 1.6 μm channels. Spatial features are similar in both retrievals and data sets, with higher values appearing mainly near the equator and in high latitudes. Differences between CLAAS-3 and MODIS vary, with higher CLAAS-3 values being more frequent. However, they remain small over large parts of the ocean. As in the all-sky LWP case, spatial features are dominated by corresponding features in ice COT (Section 7.4.5.1).

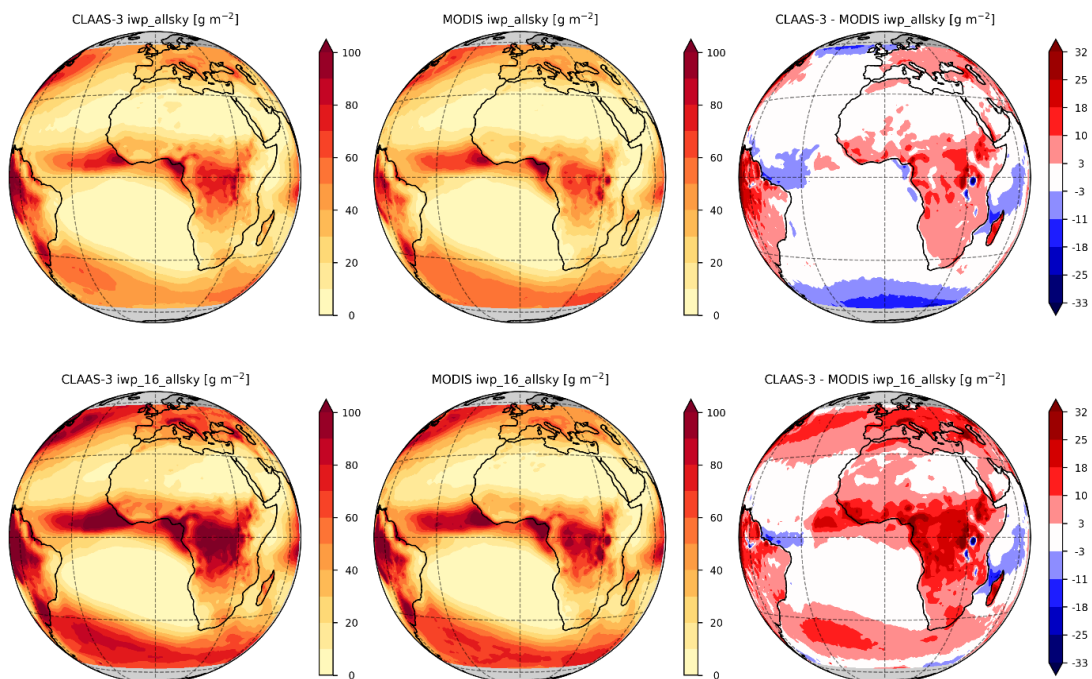


Figure 7-28: Maps of average all-sky Ice Water Path (IWP) retrieved based on the 3.9 μm channel (top row) and on the 1.6 μm channel (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left column), MODIS (middle column) and their difference (right column).

Time series comparisons of all-sky IWP from both channel pair retrievals show very good agreement in both seasonality characteristics and bias (Figure 7-29), with the latter in the 3.9

μm retrieval being remarkably small. Bias trends are also small, with some bias fluctuations in the 1.6 μm case, which could be attributed to both CLAAS-3 and MODIS variations.

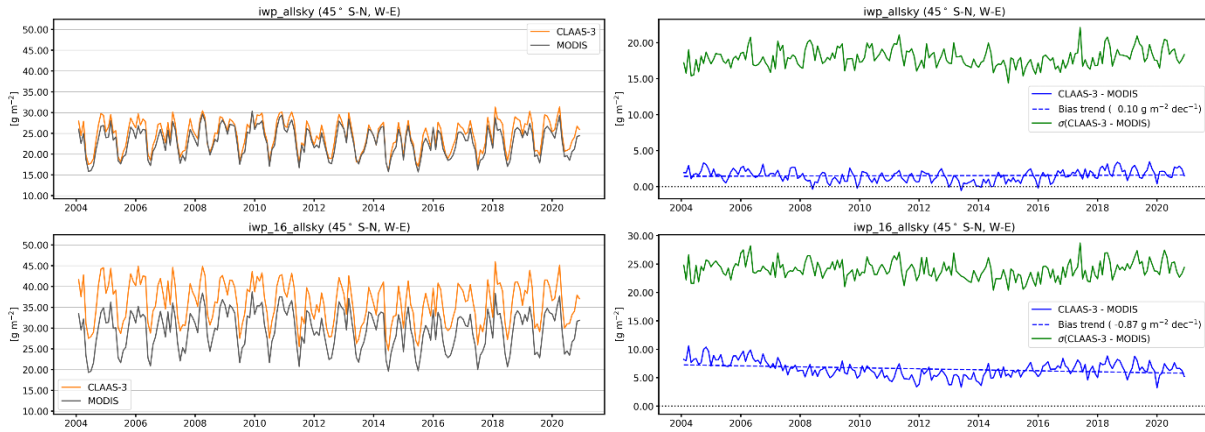


Figure 7-29: Time series of all-sky IWP retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (g m^{-2} per decade) based on linear regression.

7.4.5.1 Ice COT and CRE

As in the liquid COT case, maps of average ice COT from CLAAS-3, MODIS and their difference are shown only for the 3.9 μm retrievals (Figure 7-30), since the 1.6 μm retrieval patterns are similar. CLAAS-3 ice COT is overall lower than MODIS, with some minor exceptions over land. The largest differences appear near the equator and at the southern edges of the disk, where ice COT values also maximize.

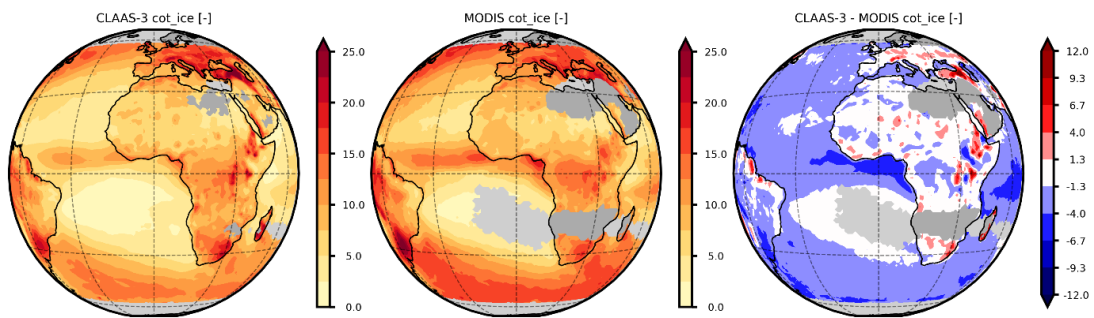


Figure 7-30: Maps of average ice Cloud Optical Thickness (COT) retrieved based on the 3.9 μm channel and based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left), MODIS (column) and their difference (right).

Ice CRE from CLAAS-3 is systematically lower than MODIS in both retrievals, with a few minor exceptions in the 1.6 μm case (Figure 7-31). Similarities in spatial patterns are also apparent.

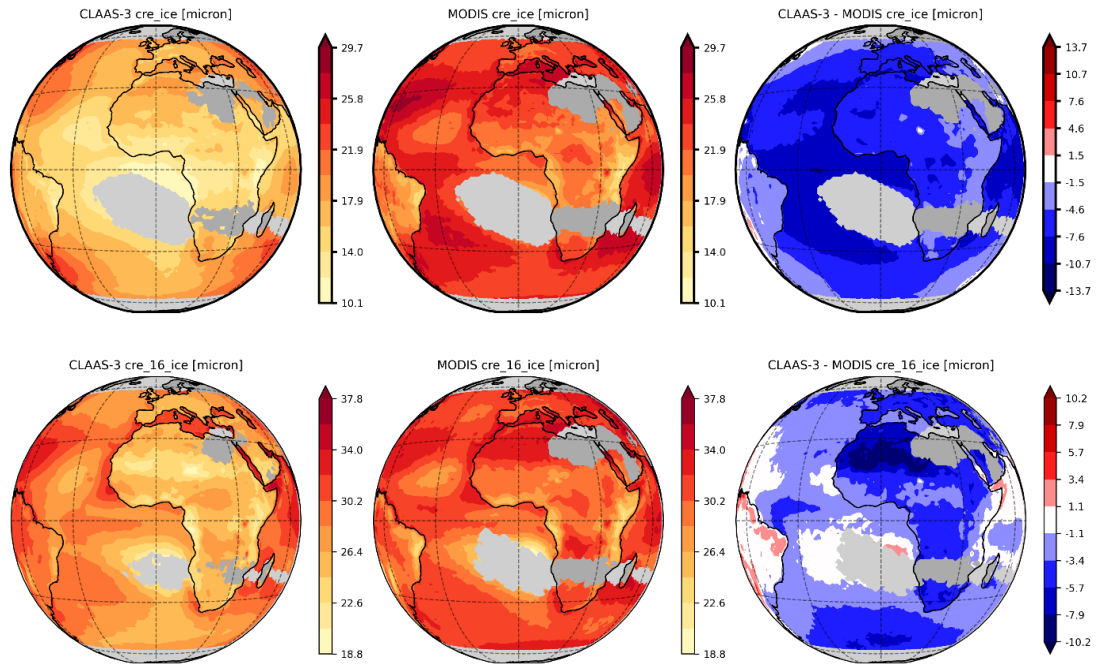


Figure 7-31: Maps of average ice Cloud Effective Radius (CRE) retrieved based on the 3.9 μm channel (top row) and on the 1.6 μm channel (bottom row) based on the entire CLAAS-3 L3 time series (2004/02 – 2020/12), from CLAAS-3 (left column), MODIS (middle column) and their difference (right column).

Time series averages of ice COT show very good agreement in both seasonality pattern and stability of bias (Figure 7-32). The bias is on average -1.0 and -1.8 in 3.9 μm and 1.6 μm retrievals, respectively. The same holds for ice CRE, with remarkably stable time series from both data sets (Figure 7-33). The average biases are -6.3 μm and -3.4 μm in the 3.9 μm and 1.6 μm retrievals, respectively.

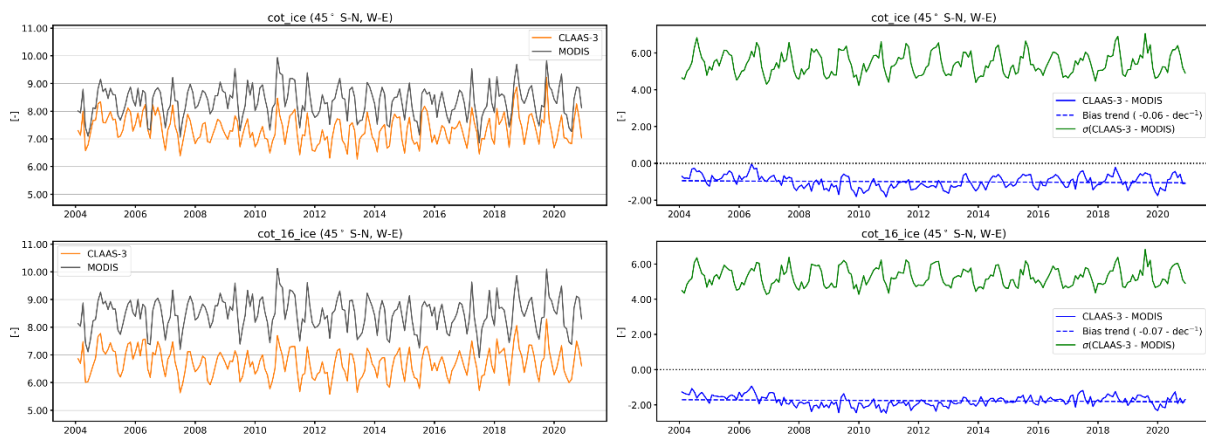


Figure 7-32: Time series of ice COT retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (per decade) based on linear regression.

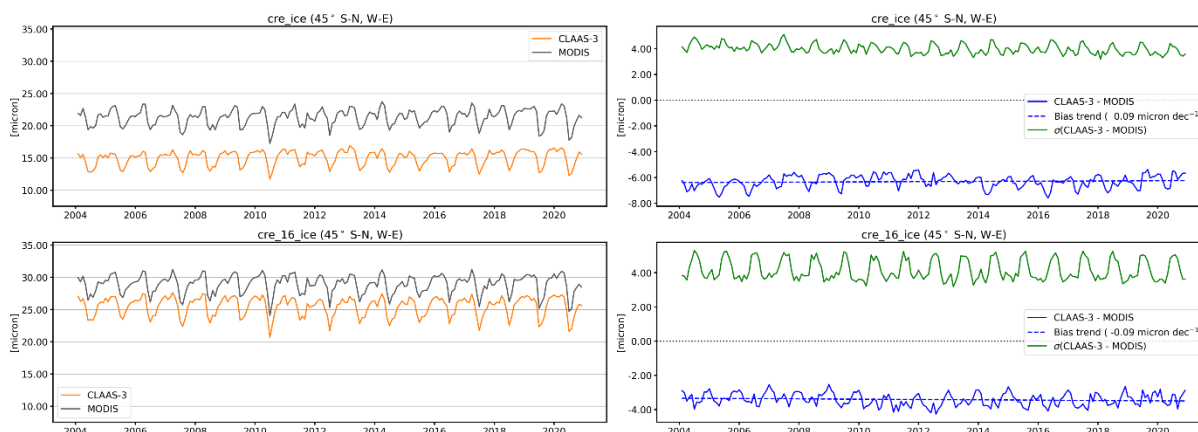


Figure 7-33: Time series of ice CRE retrieved based on the 3.9 μm channel (top row) and the 1.6 μm channel (bottom row) from monthly L3 CLAAS-3 and MODIS data (left column) and corresponding biases and bc-rmsd (right column). Time series values are averages of an area expanding from 45° S to 45° N, and from 45° W to 45° E. Blue dotted lines in the bias plots show the change in bias (μm per decade) based on linear regression.

7.4.6 JCH: Joint Cloud property Histograms of CTP and COT

For a meaningful comparison between CLAAS-3 and MODIS Joint Cloud Histograms (JCHs) of CTP and COT, common histogram bin borders are required. For this purpose, we use the “heritage” ISCCP bin borders, which provide seven bins for CTP and eight for COT. These ISCCP-like JCHs are available in MODIS L3 separately for cloudy and partly cloudy (PCL) grid cells (parameters `Cloud_Optical_Thickness_ISCCP_JHisto_vs_Pressure` and `Cloud_Optical_Thickness_PCL_ISCCP_JHisto_vs_Pressure`, respectively). We compare CLAAS-3 with both the cloudy-pixel counts only, and the sum of cloudy and partly cloudy counts. While the CLAAS-3 JCH has a finer resolution (15 bins for CTP and 13 for COT), aggregation of counts to an ISCCP-like JCH is straightforward, since the same bin borders are included in the CLAAS-3 binning system. Each ISCCP JCH bin typically corresponds to 4 CLAAS-3 JCH bins. Liquid and ice clouds are examined together in this comparison.

The CLAAS-3 and MODIS JCHs shown in Figure 7-34 include normalized values of all counts from the period 2004-2020 and the region 45°S-45°N, 45°E-45°W. For MODIS these counts are the sum of Terra and Aqua counts. The middle JCH shows counts from MODIS cloudy grid cells only, while the rightmost JCH includes counts from PCL grid cells.

The data sets are quite similar, with a prevailing low cloud and a secondary high cloud regime apparent in both. Low clouds (with CTP higher than 800 hPa and COT between 1.3 and 23) are the most frequent, constituting a very similar percentage of all cases in both data sets (22.14% in CLAAS-3 and 23.18% in MODIS). This number increases to 34.92% when MODIS PCL is included. This is because PCL counts mainly originate in (low) cloud edges, where the retrieved COT is low. The increase in occurrence of cases with low cloud COT between 0.3 and 3.6 is apparent in the rightmost JCH of Figure 7-34.

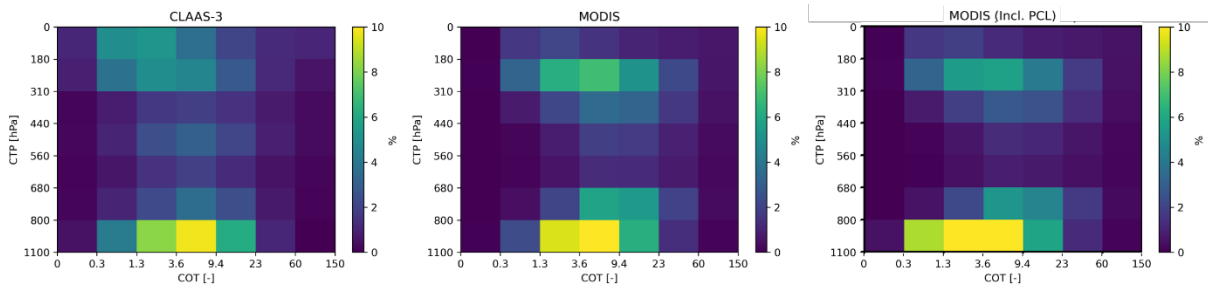


Figure 7-34: Joint Cloud Histograms (JCH) of CTP and COT from CLAAS-3 (left) and MODIS (middle, sum of Terra and Aqua counts), using the ISCCP heritage histogram bins, based on the entire CLAAS-3 time series (2004/01 – 2020/12). The rightmost JCH is also from MODIS, but includes partly cloudy (PCL) retrievals. All JCHs are normalized to 100%.

7.4.7 Summary of results

Table 7-5 summarizes the scores (bias, bc-rmsd and stability of bias) in all CLAAS-3 L3 products that were compared with MODIS. Colored cells correspond to parameters where relevant requirements are applicable. In general, threshold requirements are fulfilled in all cases except for the cloud top products bias. As shown before, however, CLAAS-3 cloud-top lies closer to CALIOP than MODIS, which is a reassuring indication of its quality. Water path results (for both liquid and ice clouds) score very well, fulfilling the optimal requirements in terms of bias and stability, and the target in terms of bc-rmsd.

Table 7-5: Time series averages of bias and bc-rmsd, and stability for all CLAAS-3 data records that were evaluated based on corresponding MODIS L3 data. Scores are color-coded where applicable: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirements.

Parameter [units]	bias	bc-rmsd	stability of bias [units/decade]
cfc [%]	-6.48	9.22	0.31
cfc_day [%]	-9.03	10.77	-0.75
cfc_night [%]	-6.51	9.83	1.31
cth [m]	2263	1625	-36.96
ctp [hPa]	-167	90	3.12
cph [%]	-5.86	13.22	2.44
cph_16 [%]	-7.13	13.05	2.54
cph_day [%]	-1.54	13.25	2.81
cph_16_day [%]	-4.89	12.98	2.99
lwp_allsky [g m ⁻²]	-3.98	10.81	0.20
lwp_16_allsky [g m ⁻²]	2.26	11.84	-0.11

Parameter [units]	bias	bc-rmsd	stability of bias [units/decade]
iwp_allsky [g m ⁻²]	1.53	17.85	0.11
iwp_16_allsky [g m ⁻²]	6.57	23.99	-0.86
cot_liq [-]	-0.86	3.28	0.19
cot_16_liq [-]	-1.44	3.51	-0.01
cre_liq [µm]	-1.89	2.03	0.07
cre_16_liq [µm]	0.89	1.85	-0.21
cdnc_liq [cm ⁻³]	14.19	41.04	-0.98
cot_ice [-]	-1.00	5.42	-0.06
cot_16_ice [-]	-1.77	5.25	-0.07
cre_ice [µm]	-6.32	3.98	0.09
cre_16_ice [µm]	-3.42	4.15	-0.09

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

8 Evaluation of CLAAS-3 ICDR

This section evaluates the consistency between the CLAAS-3 TCDR (2004-2020) and the CLAAS-3 ICDR (2021 onwards). For this purpose a backwards extension of the ICDR (covering the year 2020) was exclusively generated to allow such a thorough consistency assessment.

8.1 L2

Figure 8-1 and Figure 8-2 show spatial maps for two selected SEVIRI time slots, namely 10.01.2020 at 12 UTC and 10.07.2020 at 12 UTC, for which the L2 products from TCDR and ICDR are compared. For all products shown only minimal differences are found, except CWP for which deviations are found in some locations with the most prominent region being Eastern Europe in January. Minor CWP deviations are also found in shallow-ocean regions (e.g., the North Sea in July) and over land surfaces. These CWP deviations are due to different surface albedo data being used in the TCDR and the ICDR. The TCDR albedo over land is taken from gap-filled MODIS data, while the ICDR albedo is prescribed from a multi-year (but day-specific) climatology. This typically gives rise to modest differences in COT-CRE retrievals, in particular for thin clouds. Over snow-covered surfaces retrieval differences can be much larger because for the ICDR the albedo is derived from numerical weather prediction (NWP) snow depth and snow albedo analyses, which can differ substantially from the direct measurements used for the TCDR, and because retrievals over bright surfaces are very sensitive to the albedo. Finally, shallow ocean albedo is taken from MODIS data in the TCDR while it is prescribed with a fixed ocean value in the ICDR, which explains the observed small CWP deviations.

Table 8-1 lists corresponding deviation statistics aggregated over the full area shown. Also, noticeable deviations exist only for CWP with biases of 0.14 g/m^2 and -0.04 g/m^2 and bc-rmsd of 16.03 g/m^2 and 8.35 g/m^2 for both SEVIRI scenes. The numbers demonstrate the high consistency between the TCDR and ICDR L2 products and, as these numbers are much smaller than those obtained from the TCDR evaluation results in the previous sections, they thus validate the ICDR L2 products.

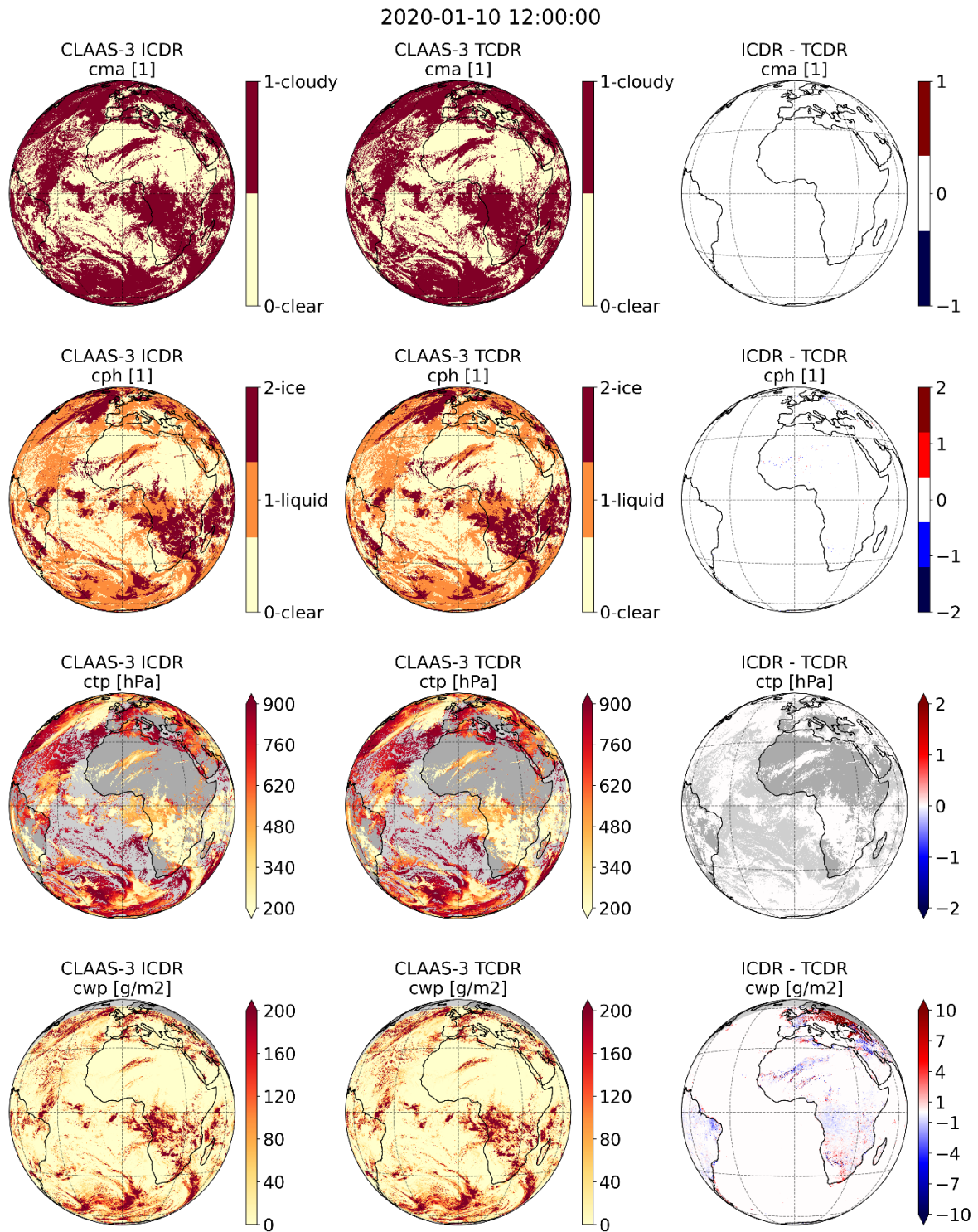


Figure 8-1: Maps of CLAAS-3 ICDR (left), CLAAS-3 TCDR (middle) and differences (right) of L2 variables on 10.01.2020 at 12 UTC.

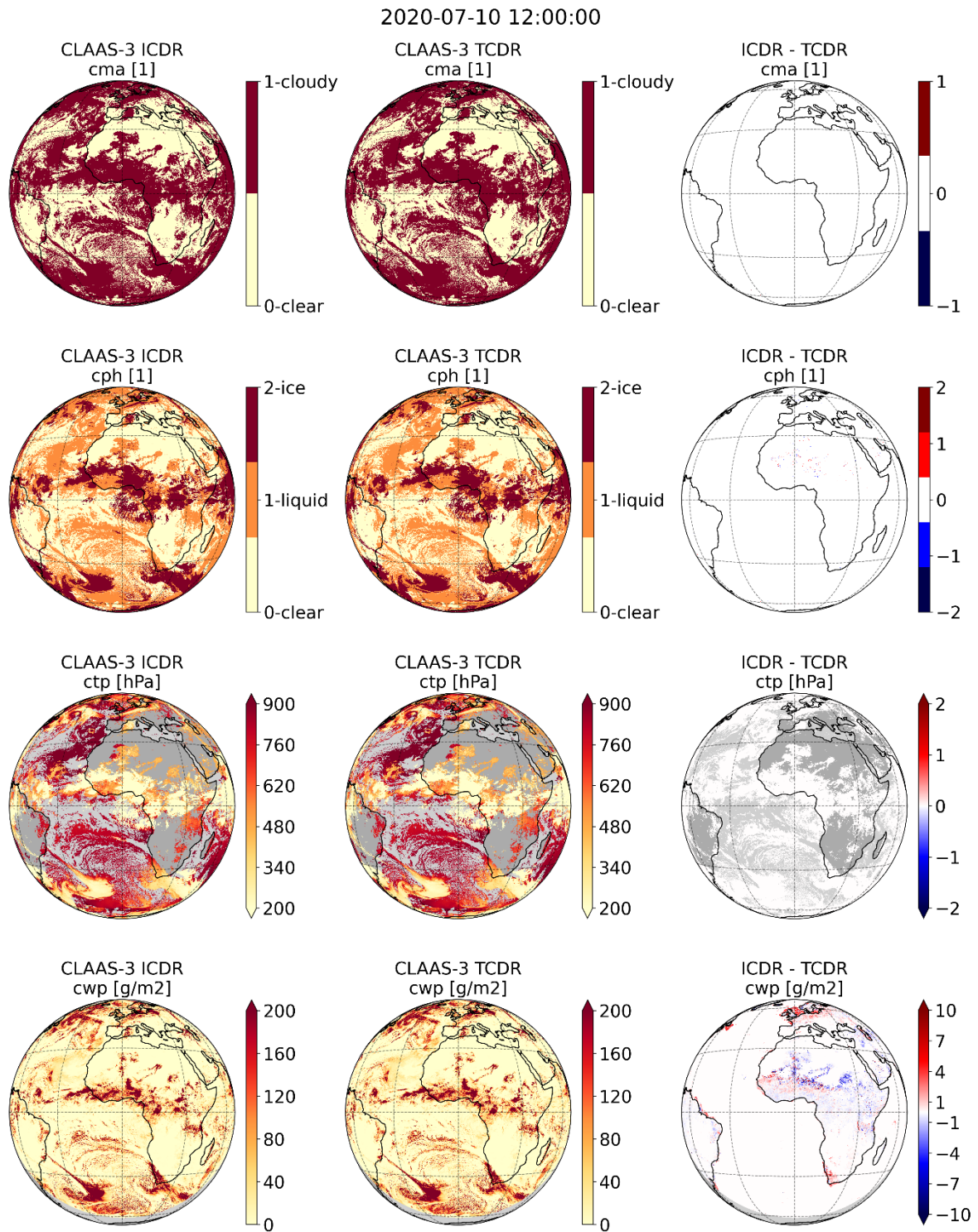


Figure 8-2: Maps of CLAAS-3 ICDR (left), CLAAS-3 TCDR (middle) and differences (right) of L2 variables on 10.07.2020 at 12 UTC.

Table 8-1: Bias and bias-corrected root mean square error for L2 cloud top pressure and cloud water path; hit rate and Hansen-Kuipers skill score for L2 cloud mask and phase for the time slots shown in figures above. Scores are color-coded where applicable: **worse than threshold**, **fulfils threshold**, **fulfils target**, and **fulfils optimal (TCDR)** requirements.

		Date	10.01.2020 12:00UTC	10.07.2020 12:00UTC
parameter	score			
ctp [hPa]	bias		0.00	0.00
	bc-rmse		0.00	0.00
cwp [g/m ²]	bias		0.14	-0.04
	bc-rmsd		16.03	8.35
cma [1]	hit rate		1.00	1.00
	KSS		1.00	1.00
cph [1]	hit rate		1.00	1.00
	KSS		1.00	1.00

8.2 L3

Figure 8-3 and Figure 8-4 show monthly mean maps for two selected months in 2020, namely January and July, for which the L3 products from TCDR and ICDR are compared. For all products shown only minimal differences are found, except LWP and IWP for which some systematic deviations are found over Eastern Europe in January. These deviations are due to different surface albedo data over snow being used in the TCDR and ICDR, as described in Section 8.1. No further deviations are noteworthy in these maps. The high agreement between TCDR and ICDR is also supported by Figure 8-5 in which time series of disk- and monthly mean values and zonal mean plots are shown including all months of 2020. The consistency is extremely high with basically no significant deviations between TCDR and ICDR.

Table 8-2 lists corresponding deviation statistics aggregated over the full spatial extent and all months in 2020. Very small numbers for bias and bc-rmsd support the findings above.

In conclusion, the assessment in this section demonstrates the high consistency between the TCDR and ICDR L3 products and thus validates the ICDR L3 products.

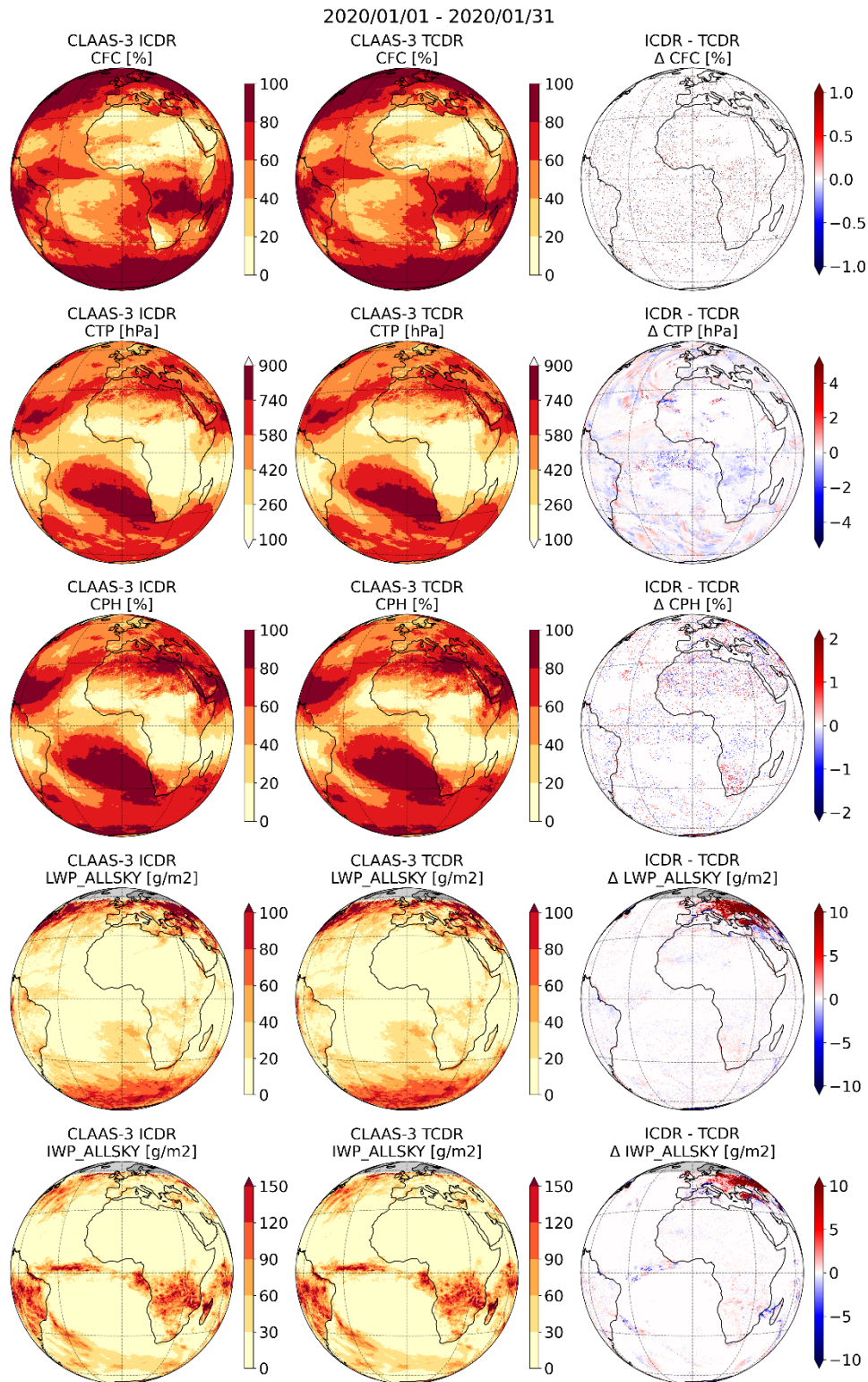


Figure 8-3: Maps of CLAAS-3 ICDR (left), CLAAS-3 TCDR (middle) and differences (right) of L3 monthly averaged products in January 2020.

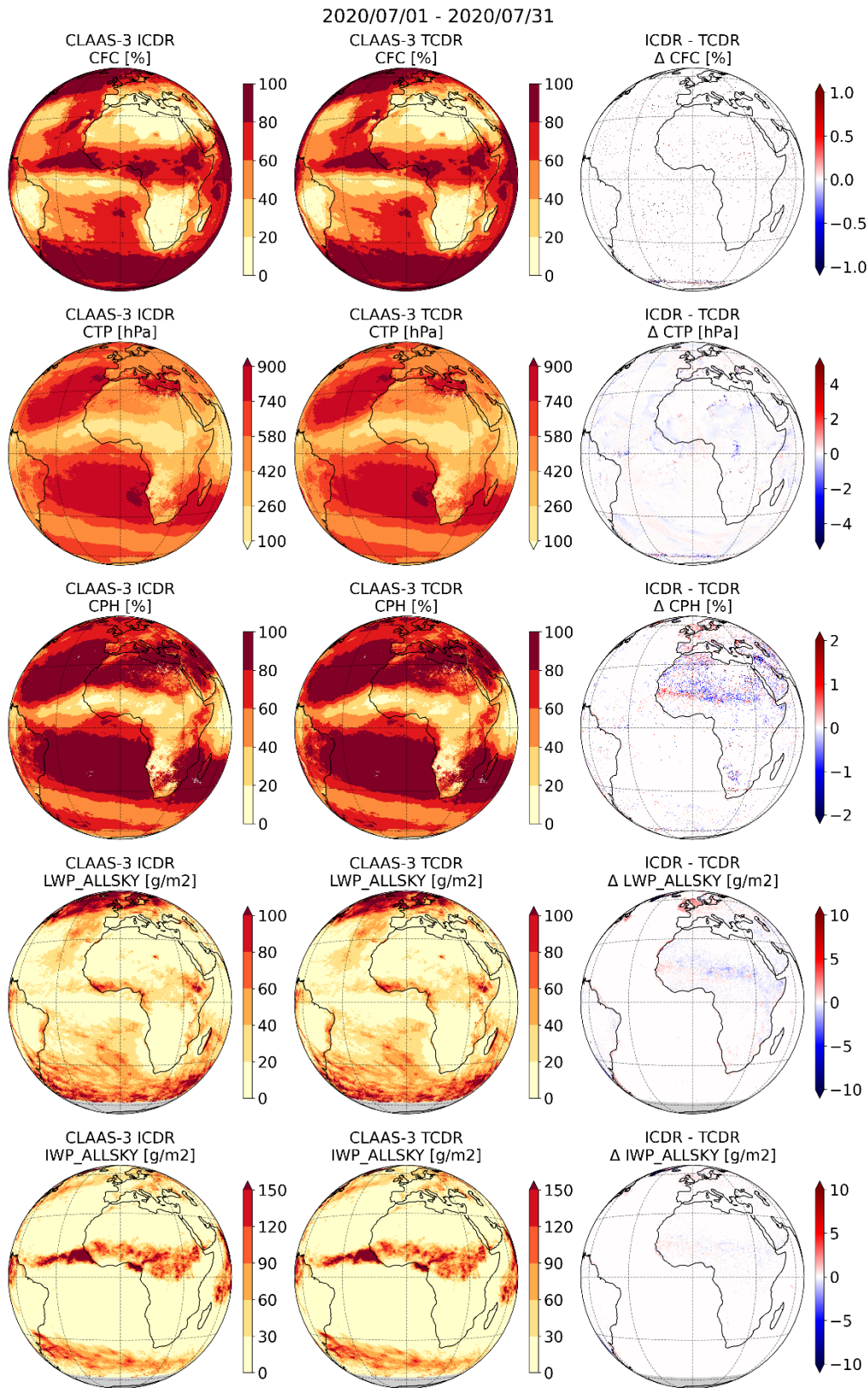
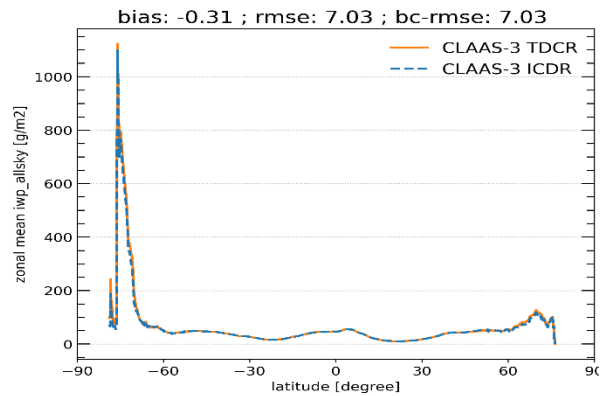
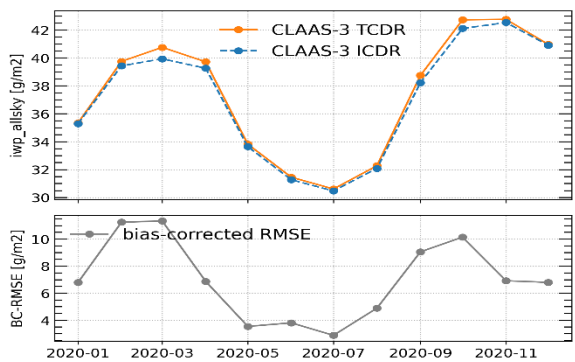
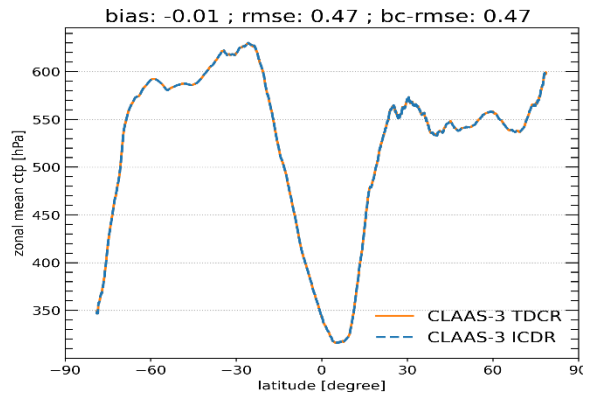
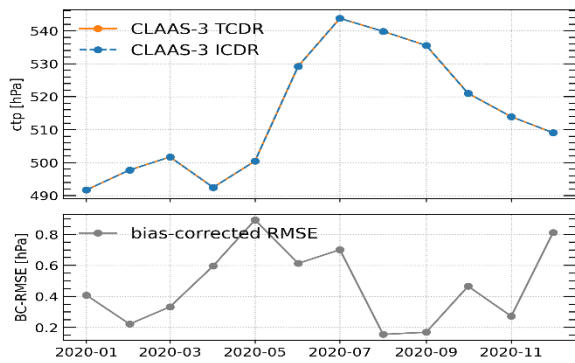
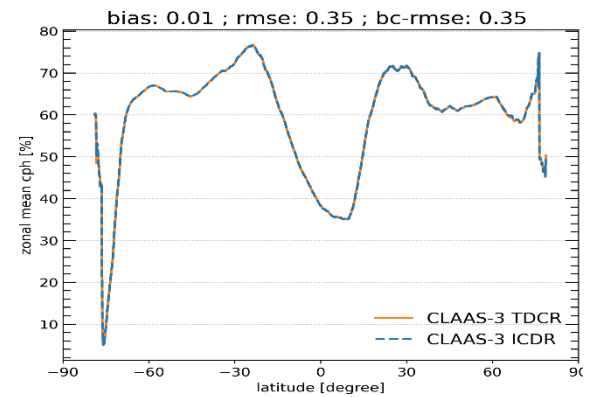
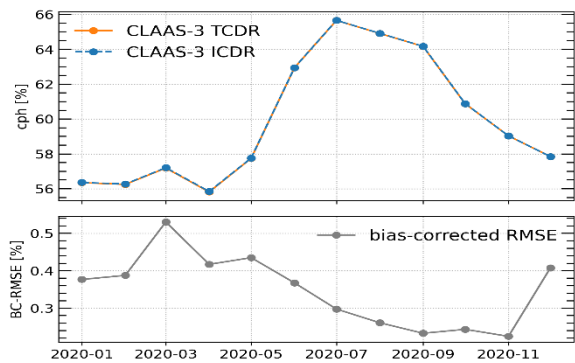
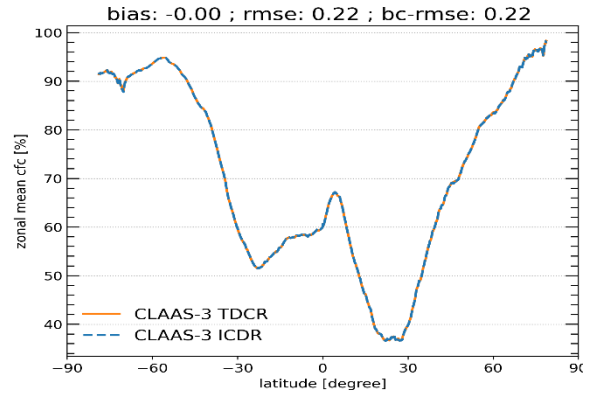
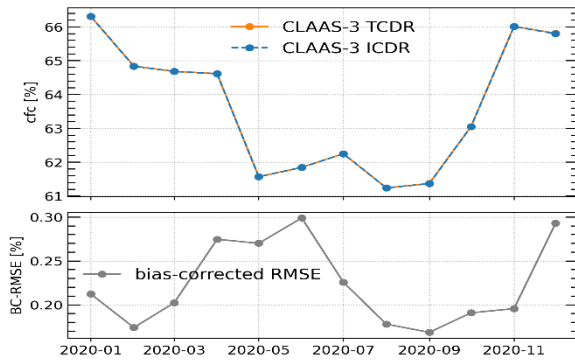


Figure 8-4: As Figure 8-3 but for July 2020.



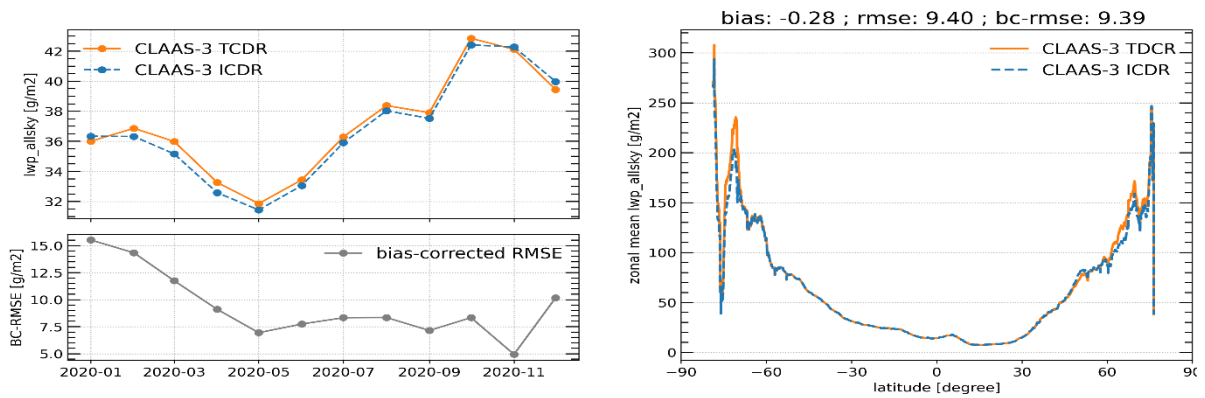


Figure 8-5: Time series and bias-corrected rmse (left), zonal mean with global averaged scores (right) of CLAAS-3 ICDR and CLAAS-3 TCDR L3 monthly averaged products for 01/01/2020-31/12/2020.

Table 8-2: Evaluation scores of CLAAS-3 TCDR and CLAAS-3 ICDR comparison. Scores are color-coded where applicable: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal (TCDR) requirements.

Parameter	Bias	bc-rmsd
cfc [%]	0.00	0.22
cph [%]	0.02	0.35
ctp [hPa]	-0.01	0.47
lwp [g/m ²]	-0.31	7.03
iwp [g/m ²]	-0.28	9.39

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

9 Conclusions

With this report we document the validation of CLAAS Edition 3 (CLAAS-3) cloud property data record. For the study we used reference datasets from independent observation sources, including different measurement strategies, as well as from similar satellite-based datasets from passive visible and infrared imagery. The used reference observation sources were MODIS, CALIOP, DARDAR, AMSR2, MAC-LWP, and SYNOP datasets. This broad spectrum of reference measurements supports the required independence and variety in the evaluation process. As such a best possible effort to assess accuracy and precision of the derived cloud properties and products has been made.

The validation was based on products at different processing levels: L2 (instantaneous data on native satellite resolution and projection) and L3 (aggregated on equal-angle latitude/longitude grid with a spatial resolution of 0.05°). Including the L2 products in the evaluation brings further insight into the precision of the derived cloud products at high temporal resolution, while the accuracy (bias) is assumed to be similar to the L3 evaluations. We evaluated the joint cloud property histogram product by inter-comparison to MODIS in order to illustrate the usefulness of this statistical approach. Table 1-1 in the executive summary gives an overview of all results with respect to the accuracy, precision, and stability requirements laid down in [AD 1]. These validation results can be summarized as follows for each individual cloud product in the CLAAS-3 TCDR:

- **Fractional Cloud Cover (CFC)**

The CLAAS-3 L2 CFC product achieves the L2 target precision requirements compared to CALIOP. Validation with SYNOP observations shows that L3 CFC meets all target requirements, and even optimal requirements for accuracy and stability. Evaluation against CALIOP L3 shows that CLAAS-3 agrees best with the CALIOP ‘top flavor’ CFC product, meeting the target requirements except for precision. Finally, L3 comparisons with MODIS demonstrate very stable results (within optimal stability requirements), but a negative mean difference outside the target range.

- **Cloud Top level (CTO)**

The CLAAS-3 CTH product yields lower mean values (slightly worse than target requirement) than the ‘cloud top’ flavor of CALIOP but agrees well with the ‘passive flavor’, for which target bias requirements are fulfilled. Comparisons of mean L2 CTH with MODIS are within the optimal requirements for liquid clouds but worse than the target for ice clouds. CLAAS-3 ice cloud CTH is closer to CALIOP than MODIS. The stability of the CTH product meets the optimal requirements in comparison to both CALIOP and MODIS. CLAAS-3 CTP evaluations are overall consistent with those for CTH.

- **Cloud Thermodynamic Phase (CPH)**

The CLAAS-3 CPH product achieves the L2 target precision requirements compared to CALIOP. Comparisons of L3 CPH with MODIS fulfil the threshold requirements. A positive trend in the monthly differences is present but appears to be mainly caused by a negative trend in MODIS liquid cloud fraction.

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

- **Liquid Water Path (LWP)**

The CLAAS-3 LWP product meets L2 target requirements for bias and bc-rmsd against AMSR2, while the L2 comparisons with MODIS fulfil the threshold requirements. The L3 validation against LWP from microwave radiometers focusing on the south-east Atlantic shows that the target requirements are fulfilled. In comparison with MODIS, LWP satisfies the target requirements for precision and optimal requirements for accuracy and stability.

- **Ice Water Path (IWP)**

The CLAAS-3 IWP L2 product does not fulfil the threshold requirements against DARDAR observations, and only threshold requirements in comparison to MODIS. This is partly a consequence of the large range of IWP values, where differences between products can become large for the higher values, leading to relatively large bias and bc-rmsd even if correlations are reasonable. On the other hand, comparisons of L3 IWP with MODIS fulfil the optimal requirements for accuracy, precision and stability.

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

This product is excluded from specific requirement testing because it is composed of three already evaluated products (CPH, COT and CTP). Nevertheless, inter-comparisons with the corresponding MODIS product show reasonable agreement. JCH provides added value to the underlying products by giving important clues on the statistical distribution of the involved parameters. It is believed that the access to this product representation will greatly enhance the usefulness of CLAAS-3 for some applications.

Apart from the validation and evaluation with external datasets, the systematic differences between CLAAS-3 and its predecessor CLAAS-2.1 were characterized. This showed that apart from the extended temporal range (three years were added), new product layers (e.g., liquid cloud droplet number concentration) and more detailed retrieval error estimates, the quality and stability of several products has improved considerably. However, a number of weaknesses were also identified, and these are planned to be tackled in the next CLAAS edition.

Until this next edition will be ready (tentatively in 2025), CLAAS-3 data will be produced in the form of an ICDR, including cloud products from 2021 onward, which are continuously replenished in present time. Exclusively for this Validation Report, a backwards-extension (covering the year 2020) was generated in order to evaluate the consistency with the TCDR. The evaluation revealed a very high consistency between the ICDR and the TCDR with differences much smaller than the deviations of the TCDR to the reference observations. The ICDR will be subject to annual quality assessments in the future.

CLAAS-3 provides a wealth of information on cloud properties at high spatial and temporal resolution. As such it can be useful for a range of applications such as NWP model evaluation and process studies. With its length, 17 years for the TCDR, and continuously growing with the ICDR, it is also becoming of increasing interest for climate studies. The detailed evaluation

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

and validation in this report provides a solid characterization of the CLAAS-3 products, which should aid users in understanding the quality and limitations of the data and thereby assessing the utility of the products for their specific applications.

10 Appendix A: SEVIRI L1.5 data

The time series of SEVIRI reflectances was carefully calibrated against the Moderate Resolution Imaging Spectroradiometer (MODIS), which improved especially the retrieval of microphysical parameters. The calibration method is outlined in Meirink et al. (2013). Because the same MODIS instrument (on the Aqua satellite) was used to calibrate all MSG satellites, the calibration also ensured homogenisation between the SEVIRI instruments. Resulting calibration slopes can be found in Figure 3-2 in [RD 2].

The IR radiances of SEVIRI were used as provided by EUMETSAT, relying on the on-board black-body calibration, since a Fundamental Climate Data Record (FCDR) is not yet available. We have analysed the stability of the IR observations by plotting brightness temperatures (BT) at midnight averaged over a large area (Figure 10-1). Variations in these time series can occur due to natural variability. However, large variations, in particular if they are not shared between different channels, can be a sign of calibration issues. Figure 10-1 gives a very strong indication that the 13.4 μm channel suffers from a spurious decrease in BT over time, especially at the transition between Meteosat-8 and -9. The CLAAS-3 algorithms do not use this channel so this has no effect on the retrieved cloud properties.

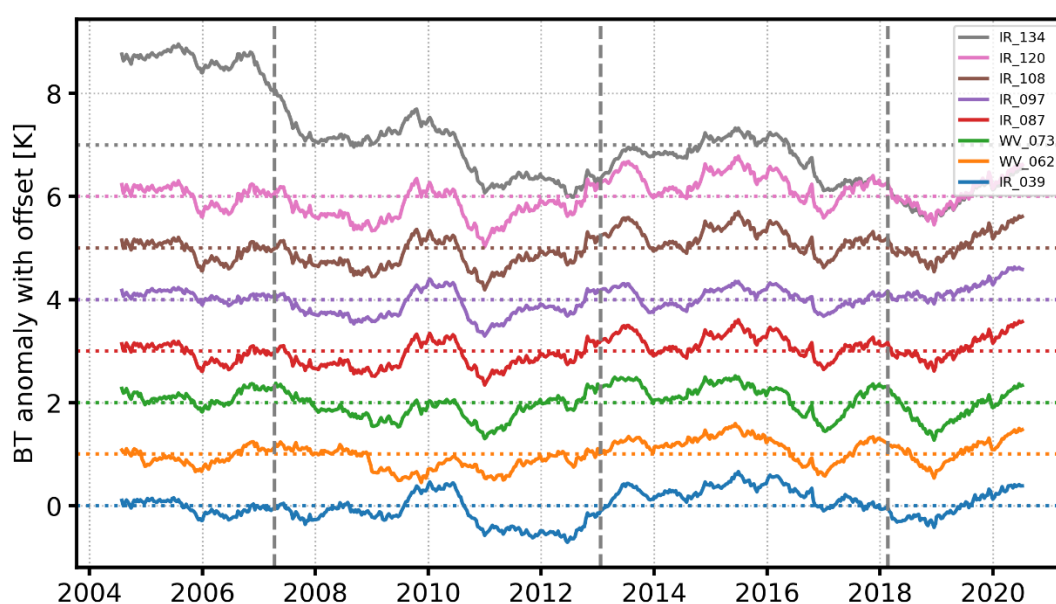


Figure 10-1: Time series of brightness temperature measured at 00:00 UTC in SEVIRI IR channels. The measurements have been averaged over a 2112 x 2112 pixel area centred around 0°/0° latitude/longitude, which corresponds to about (30 °W – 30 °E , 30 °S – 30 °N). The solid lines show 1-year running mean BT, consecutively offset by 1 K, as indicated by the corresponding dotted lines. Vertical, dashed grey lines indicate the transitions between the prime MSG satellites.

The other channels are much more similar. However, differences between channels, as displayed in Figure 10-2, do show irregularities on longer time scales. An example is the difference between 3.9 μm and 10.8 μm BT. While this difference is generally stable at a value of around -0.2 K, it is about 0.5 K lower (more negative) in the years 2011-2012. In the latest years of the record, it also appears to be slightly lower. Both deviations might be a sign of calibration issues. The 3.9 – 10.8 μm BT difference is an important feature in the probabilistic

cloud mask algorithm, and calibration issues can thus result in cloud cover anomalies, especially during nighttime, when no shortwave measurements are available to provide additional information on cloud cover.

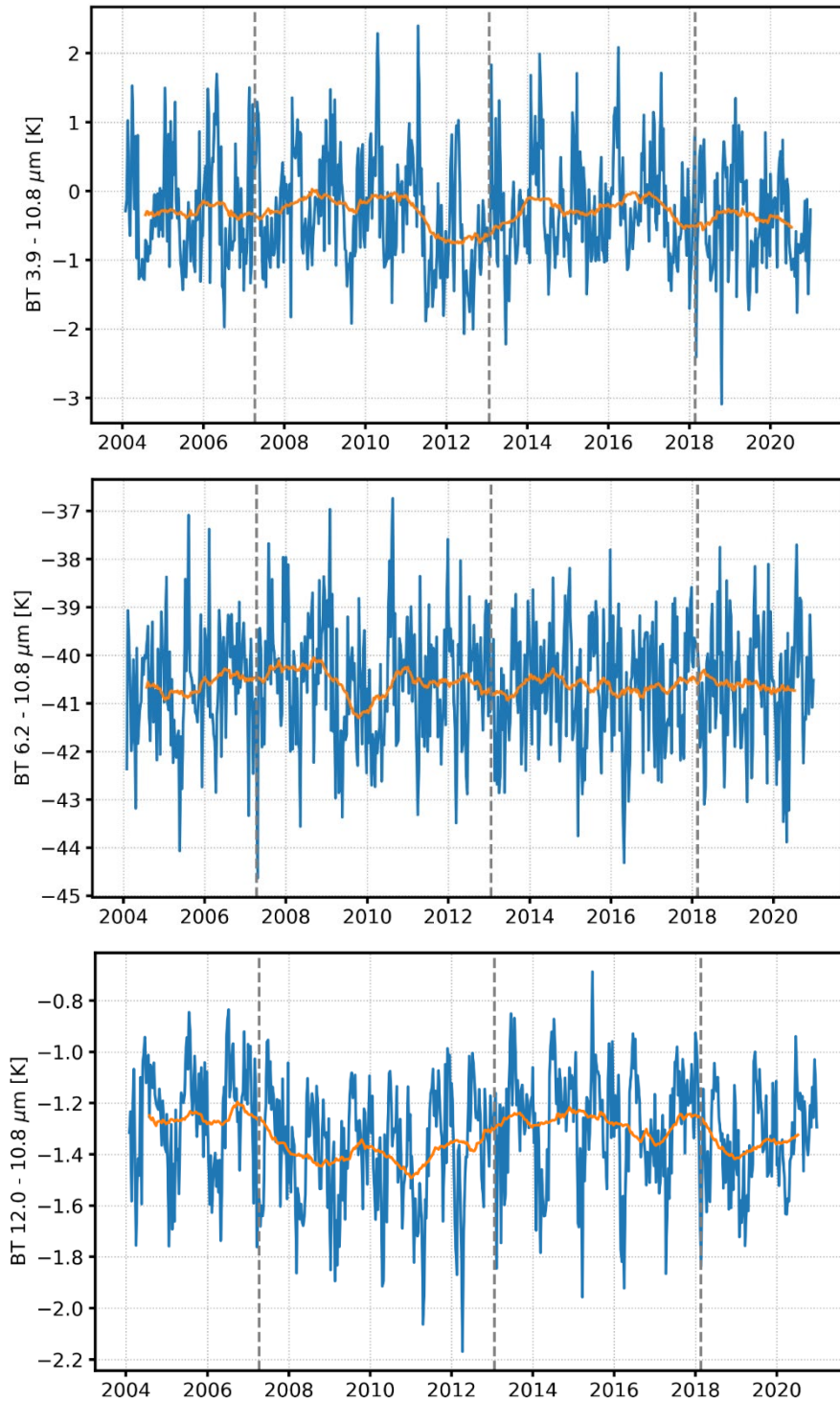


Figure 10-2: Difference between SEVIRI IR channels: BT 3.9 – BT 10.8 μm (top), BT 6.2 – BT 10.8 μm (middle), BT 12.0 – BT 10.8 μm (bottom). Blue lines show weekly values, while the orange lines are 1-year running means. Data have been spatially averaged as outlined in Figure 10-1. Vertical, dashed grey lines indicate the transitions between the prime MSG satellites.

	<p style="text-align: center;">Validation Report SEVIRI cloud products CLAAS Edition 3</p>	<p>Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022</p>
---	---	---

The 6.2 μm water vapour absorption channel appears to be very stable with respect to the 10.8 μm window channel, although a possible anomaly is present near the end of 2009. Finally, differences between the 12.0 and 10.8 μm window channels are generally very small and (again) rather stable. However, in the years during which the MSG-2 and MSG-4 SEVIRI instruments were active, i.e. 2007-2012 and 2018-2020, respectively, slightly larger differences are observed than in the other years. This might point to calibration issues, with possible effects on cloud property retrievals.

The CLAAS-3 TCDR is complemented by an ICDR, which has at this stage been produced for the year 2020 with exactly the same SEVIRI L1.5 data as for the TCDR. For the extension of the ICDR into present, the shortwave calibration as outlined in Figure 3-2 in [RD 2] will be used, i.e. the temporal trends in calibration coefficients determined with past data will be retained. For the IR channels the operational, on-board calibration will be used, as has been done for the TCDR.

For a more detailed instrument specification and description of the calibration the reader is referred to the SEVIRI Algorithm Theoretical Baseline Document [RD 2] and the SEVIRI Product User Manual [RD 1].

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

11 Appendix B: Reference datasets

In this Appendix the datasets used as reference for the validation and evaluation of CLAAS-3 are described in detail.

11.1 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 datasets used to evaluate the cloud fractional coverage estimates. At manned stations the total cloud cover is visually estimated by human observers. In contrast, ceilometers are used for that purpose at automatic stations. However, for data quality and consistency reasons, only those SYNOP reports provided by manned airport stations were taken into account (~1800 stations globally).

SYNOP cloud cover observations are used for the evaluation of the L3 CFC product.

Manual cloud observations are affected by many sources of error. We list some of the most important ones 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, 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. This is natural 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 become 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 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.

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

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 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 nighttime. However, the variability (precision) probably reaches higher absolute values and it is largest during night conditions. This may lead to a strong seasonal variation in quality 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 nighttime but there is also a considerable risk that the precision figures degrade, mainly because ceilometers only observe 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, particularly due to their long-term availability.

11.2 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 (assuming a successful aerosol-cloud discrimination) 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 11-1).

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 June 2006. CALIOP provides detailed profile information about cloud and aerosol particles and corresponding physical parameters.

CALIOP measures the backscatter intensity at 1064 nm while two other channels measure the orthogonally polarized components of the backscattered signal at 532 nm. 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 measurement is 333 m and the vertical resolution is 30-60 m.

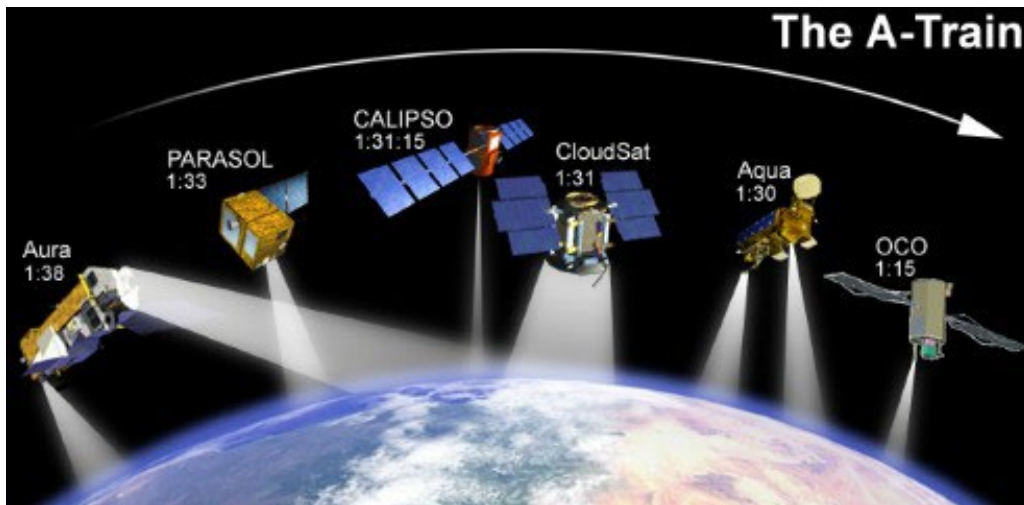


Figure 11-1: The Aqua-Train satellites. (Image credit: NASA)

The CALIOP products are available in three different versions with respect to the along-track resolution ranging from 333 m (resolution based on the spacing between consecutive footprints of 70 m), 1 km and 5 km. The two latter resolutions are consequently constructed from several original footprints/FOVs (often called “single shots”). Of importance is also that, for the 5 km product, thin clouds have been identified after integrating over even larger distances (20 km and 80 km) to maximize the signal-to-noise ratio. This allows a higher confidence in the detection and identification of cloud and aerosol layers compared to the original footprint resolution. For example, the identification of very thin Cirrus clouds is more reliable in the 5 km dataset than in the 1 km dataset since signal-to-noise levels have been raised by using a combined dataset of several original profiles.

We used the CALIOP L2 5km cloud layer (CLAY) dataset version 4.20 (CALIPSO Science Team, 2021) for validation purpose since this resolution is closest to the nominal SEVIRI resolution and because of the advantages described above. It 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.

CALIOP data is used for L2 validation of the CFC, CTH and CPH products. It has the following uncertainties and error sources.

It should be emphasized that the CALIOP measurement is probing the atmosphere very efficiently in the along-track direction since it is a near-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 70 m, the individual footprint of the lidar beam. Thus, to compare CALIOP-derived results with the results of 3 km SEVIRI pixel data is not entirely consistent (i.e., CALIOP is only capable of covering the SEVIRI 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 SEVIRI 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 SEVIRI pixel that coincides with a CALIOP measurement. Most clouds will have aspect ratios for the two horizontal directions that guarantee detection by CALIOP.

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

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 by 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 (i.e., saying something on the performance of the applied cloud detection method for clouds that should be theoretically detectable) one should consider filtering out the contributions from the very thinnest clouds. We have tested this approach in this validation study, both in the study of cloud amount (CFC), cloud phase (CPH) and cloud top height (CTH).

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 nighttime and daytime data.

In the CLAAS-3 validation context we have used the 5 km CALIOP dataset (product CLAY versions 4.20) since this resolution is closest to the nominal SEVIRI resolution. However, the results in different datasets from CALIPSO, related to different horizontal resolutions (with the three options 333 m, 1 km and 5 km) are unfortunately not entirely consistent (e.g., Karlsson and Johansson, 2013). It means that some of the thick (opaque) boundary layer clouds that are reported in fine resolution (333 m and 1 km) datasets are not reported in the higher resolution (5 km or higher) datasets. This has to do with the methodology to do averaging at the longer scales (5 km or higher) where contributions from strongly reflecting boundary layer clouds are removed from the original signal to facilitate detection of very thin cloud layers and aerosols. To correct for this inconsistency, a correction is applied to the 5 km CLAY product meaning that a 5 km FOV, which is declared cloud-free in the official CLAY product, is reclassified as cloudy if a majority of 333 m FOVs within the 5 km FOV are cloudy. The tool used for enabling matchups between SEVIRI and CALIOP and for applying these corrections to CALIOP products is called *atrain-match* (see https://github.com/foua-pps/atrain_match).

However, for the evaluation of L3 products, we have used a special L3 product based on CALIOP data prepared for the [GEWEX cloud assessment study](#). [This dataset](#) has a horizontal resolution of 1° and is based on results from the L2 CLAY product version 4.20. It includes monthly averaged cloud parameters in four, so called, ‘flavors’ developed following different philosophies. In this validation study we used top layer and passive flavors. The top layer flavor is based on the top layer cloud only in each profile. Passive flavor chooses the top layer cloud which would be detected by a passive sensor, typically choosing the layer with the optical depth greater than or equal to 0.3.

Some differences between this product and experimental L3 compilations based on the corrected CLAY 5 km dataset (mainly used to evaluate L2 products) are noticed. Especially, the GEWEX product gives somewhat lower cloud amounts over regions with predominantly boundary layer cloudiness (e.g. marine stratocumulus areas). This difference is interpreted as coming from an attempt in the GEWEX product to reduce the potential impact of overestimated amounts of thin water clouds detected using large integration scales (i.e., above 5 km).

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

In conclusion: despite the fact that the CALIPSO cloud observations most likely are the best available cloud reference dataset being released so far, we might still see a negative bias of a few percent when comparing to the CALIOP-derived cloud cover when using the 5 km dataset because of the CALIOP contribution from clouds being subvisible to passive sensors. 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 accumulated results based on a large number of full global orbits. This also concerns problems with reduced signal-to-noise ratios due to solar contamination during daytime.

11.3 DARDAR

To complement the picture drawn by the CALIOP lidar also the Cloud Profiling Radar (CPR) onboard CloudSat, launched in 2006, is considered. CPR is a nadir-looking radar sensing the atmosphere at 94 GHz. The instrument sensitivity is defined by a minimum detectable reflectivity factor of -26 dBZ and the calibration accuracy is better than 1.5 dB.

The DARDAR dataset (Cazenave et al., 2019; 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 (\sim 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 between these different regimes. The DARDAR product has the vertical resolution of CALIOP (60 m) and a horizontal resolution given by the radar footprint (1.1 km). For the current evaluation, DARDAR v3 was used, which features several improvements compared to v2 (Cazenave et al., 2019). DARDAR v3 data were downloaded from the ICARE FTP website: <ftp://ftp.icare.univ-lille1.fr/>. Two flavors are available: v3.00 using a combination of mass-size relationships from Brown and Francis (1995) and Mitchell (1996), and v3.10 using composite mass-size relationships from Heymsfield et al. (2010).

DARDAR data is used for L2 evaluation of ice cloud optical thickness, ice particle effective radius and IWP. It should be noted that identical results are not expected due to the very different measurement principles between the respective datasets. For example, comparisons 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.

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

11.4 MODIS

The Moderate Resolution Imaging Spectroradiometer (MODIS) is an advanced imaging instrument onboard the Terra (EOS AM, launched in 1999) and Aqua (EOS PM, launched in 2002) polar satellites (see <http://modis-atmos.gsfc.nasa.gov/index.html>). Terra's orbit around the Earth is sun synchronous and timed so that it 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.

Because the MODIS instruments are among the most advanced passive imagers in space, have proven to be very stable over time, and have been flying during the complete MSG time frame, the cloud products from MODIS Collection 6.1 (Platnick et al., 2017) are used here as a reference.

Both L2 and L3 products have been used. MODIS L2 cloud products (MOD06_L2 for Terra and MYD06_L2 for Aqua) are available in 5-minute granules containing a range of retrieved cloud properties and ancillary data on 1 km and 5 km spatial resolutions. Several cloud properties at 1 km resolution have been considered for the evaluations presented in Section 6.4. The variables CDNC and CGT are not included in the MODIS L2 products. We have derived them from the COT and CRE (based on the 3.7 μm channel) measurements following the approach of Bennartz and Rausch (2017), which was also used to derive CLAAS-3 CDNC and CGT.

L3 gridded atmosphere monthly global products (MOD08_M3 for Terra and MYD08_M3 for Aqua) are used for the comparisons in Section 7.4. They contain monthly global 1×1 degree grid average values of atmospheric parameters related to atmospheric aerosol particle properties, total ozone burden, atmospheric water vapour, cloud properties, and atmospheric stability indices. Monthly CDNC and CGT are obtained from the data record described in Bennartz and Rausch (2017).

Uncertainties in MODIS retrievals are in general expected to be somewhat smaller than what can be obtained with SEVIRI retrievals, because MODIS has a wider variety of channels. For example: multiple CO_2 channels in principle give additional cloud-top height information, 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 CLAAS-3. Therefore, the inter-comparisons have to be viewed as an evaluation rather than pure validation, allowing to judge the overall consistency of the CLAAS-3 data record.

11.5 Microwave imagers

Passive microwave imagers, such as the Advanced Microwave Scanning Radiometer – EOS (AMSR-E), can be used to retrieve column-integrated cloud liquid water over ocean along with column water vapour, sea surface temperature, 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

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

rain water path has to be made and subtracted from the total liquid water path to retrieve the cloud liquid water path.

Microwave imagers are used for both L2 and L3 validation of CLAAS-3 LWP. For L2 comparisons data from AMSR2 are used. AMSR2 is flown on the Global Change Observation Mission for Water (GCOM-W) satellite launched by the Japan Aerospace Exploration Agency (JAXA) in 2012. GCOM-W is in a sun-synchronous orbit with an ascending node at 13:30 equatorial overpass time. AMSR2 is a dual-polarization conical-scanning passive microwave radiometer with 16 channels ranging from 6.9 to 89 GHz. The instrument was designed to measure ice, snow, water vapour, precipitation, cloud properties, sea surface temperature and wind speed. Here the daily Air-Sea Essential Climate Variables product version 8.2 generated by Remote Sensing Systems (RSS) and available at <https://www.remss.com/missions/amr>, is used (see Hilburn and Wentz (2008) and references therein for more information on the algorithms). The data are spatially aggregated to a global equal-angle 0.25 x 0.25 degree grid, separately for ascending and descending nodes, of which the (daytime) ascending node is used for this evaluation. AMSR2 LWP products are derived primarily from the liquid sensitive 37 GHz channel measurement, which has a spatial footprint of 7 x 12 km. Thus, in the gridded product typically eight or less L2 retrievals are averaged.

For L3 comparisons, the Multisensor Advanced Climatology of Liquid Water Path (MAC-LWP; Elsaesser et al., 2017) version 1 was chosen as an independent reference dataset. MAC-LWP v1 is based on version 7 RSS retrievals from various microwave radiometer instruments, including the Special Sensor Microwave Imager (SSM/I) series, the Tropical Rainfall Measurement Mission Microwave Imager (TMI), the Global Precipitation Measurement Microwave Imager (GMI), AMSR-E, and AMSR2. The data record contains monthly 1 x 1 degree gridded oceanic LWP information spanning the years 1988 – 2016. Both monthly means and (average) monthly mean diurnal cycles are provided. The climatology is corrected for drifting satellite overpass times and for a clear-sky bias known to exist in the LWP retrievals. Statistical error estimates are provided, ranging between about 10% and 25% depending on location (Elsaesser et al., 2017). However, apart from the clear-sky bias, for which a correction has been applied in MAC-LWP, various other systematic errors are present in MW LWP retrievals, including cloud-rain partition bias, cloud temperature bias, and cloud-fraction-dependent bias (Greenwald et al., 2018), which are not included in the error estimates.

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 grid cells without ice clouds (in L2) or with a sufficiently low monthly mean ice cloud fraction (in L3). In the MSG disk the requirement for L3 leads to selecting an area in the southeast Atlantic off the Namibian coast with persistent marine stratocumulus (Sc). Specifically, the region from 20° – 10°S and 0° – 10°E was used for LWP evaluation.

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

12 References

Baum, B. A., P. Yang, A. J. Heymsfield, C. G. Schmitt, Y. Xie, A. Bansemmer, Y.-X. Hu, and Z. Zhang, 2011: Improvements in shortwave bulk scattering and absorption models for the remote sensing of ice clouds, *J. Appl. Meteorol. Clim.*, 50, 1037-1056, doi:10.1175/2010JAMC2608.1.

Benas, N., Finkensieper, S., Stengel, M., van Zadelhoff, G.-J., Hanschmann, T., Hollmann, R., and Meirink, J. F., 2017: The MSG-SEVIRI-based cloud property data record CLAAS-2, *Earth System Science Data*, 9, 415–434, doi:10.5194/essd-9-415-2017.

Bennartz, R. and J. Rausch, 2017: Global and regional estimates of warm cloud droplet number concentration based on 13 years of AQUA-MODIS observations, *Atmos. Chem. Phys.*, 17, 9815-9836, doi:10.5194/acp-17-9815-2017.

Brown, P. R. A. and Francis, P. N., 1995: Improved measurements of the ice water content in cirrus using a total-water probe, *J. Atmos. Ocean. Tech.*, 12, 410–414.

CALIPSO Science Team, 2021: CALIPSO/CALIOP Level 2, Lidar 5 km Cloud Layer Product, version 4.20, Hampton, VA, USA: NASA Atmospheric Science Data Center (ASDC). Accessed 2021-11-26 at 10.5067/CALIOP/CALIPSO/LID_L2_05KMCLAY-STANDARD-V4-20.

Cazenave, Q., M. Ceccaldi, J. Delanoë, J. Pelon, S. Groß, and A. Heymsfield, 2019: Evolution of DARDAR-CLOUD ice cloud retrievals: new parameters and impacts on the retrieved microphysical properties, *Atmos. Meas. Tech.*, 12, 2819–2835, doi:10.5194/amt-12-2819-2019.

Chepfer H., S. Bony, D. M. Winker, G. Cesana, J.L. Dufresne, P. Minnis, C.J. Stubenrauch, S. Zeng, 2010: The GCM Oriented CALIPSO Cloud Product (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.

Elsaesser, G.S., O'Dell, C.W., Lebsock, M.D., Bennartz, R., Greenwald, T.J., and Wentz, F.J., 2017: The Multisensor Advanced Climatology of Liquid Water Path (MAC-LWP). *J. Climate*, 30, 10193–10210, doi:10.1175/JCLI-D-16-0902.1.

GCOS, 2011: Systematic observation requirements for satellite-based products for climate <http://www.wmo.int/pages/prog/gcos/Publications/gcos-154.pdf>.

Greenwald, T. J., 2009: A 2 year comparison of AMSR-E and MODIS cloud liquid water path observations. *Geophys. Res. Letters*, 36, L20805. <https://doi.org/10.1029/2009GL040394>.

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

Greenwald, T. J., Bennartz, R., Lebsock, M., & Teixeira, J., 2018: An Uncertainty Data Set for Passive Microwave Satellite Observations of Warm Cloud Liquid Water Path. *J. Geophys. Res.: Atmospheres*, 123, 3668–3687. <https://doi.org/10.1002/2017JD027638>.

Håkansson, N., Adok, C., Thoss, A., Scheirer, R., and Hörnquist, S.: Neural network cloud top pressure and height for MODIS, *Atmos. Meas. Tech.*, 11, 3177–3196, <https://doi.org/10.5194/amt-11-3177-2018>, 2018.

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

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

Hilburn, K. A. and F. J. Wentz, 2008: Intercalibrated passive microwave rain products from the Unified Microwave Ocean Retrieval Algorithm (UMORA), *Journal of Climatology and Applied Meteorology*, 47, 778–794.

Heymsfield, A. J., Schmitt, C., Bansemer, A., and Twohy, C. H., 2010: Improved Representation of Ice Particle Masses Based on Observations in Natural Clouds, *J. Atmos. Sci.*, 67, 3303–3318, <https://doi.org/10.1175/2010JAS3507.1>.

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. and E. Johansson, 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, *Atmos. Meas. Tech.*, 6, 1271–1286, doi:10.5194/amt-6-1271-2013.

Karlsson, K.-G. and Håkansson, N.: Characterization of AVHRR global cloud detection sensitivity based on CALIPSO-CALIOP cloud optical thickness information: demonstration of results based on the CM SAF CLARA-A2 climate data record, *Atmos. Meas. Tech.*, 11, 633–649, <https://doi.org/10.5194/amt-11-633-2018>, 2018.

Meirink, J.F., R.A. Roebeling and P. Stammes, 2013: Inter-calibration of polar imager solar channels using SEVIRI, *Atm. Meas. Tech.*, 6, 2495–2508, doi:10.5194/amt-6-2495-2013.

Mitchell, D., 1996: Use of mass- and area-dimensional power laws for determining precipitation particle terminal velocity., *J. Atmos. Sci*, 53, 1710–1723.

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.

Platnick, S., 2000: Vertical photon transport in cloud remote sensing problems. *J. Geophys. Res.: Atmospheres*, 105, 22919–22935, doi:10.1029/2000jd900333.

Platnick, S., Meyer, K. G., D., K. M., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L., and Riedi, J., 2017: The

	Validation Report SEVIRI cloud products CLAAS Edition 3	Doc. No: SAF/CM/KNMI/VAL/SEV/CLD Issue: 3.1 Date: 08.08.2022
---	--	--

MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua, *IEEE T. Geosci. Remote*, 55, 502–525, doi: 10.1109/TGRS.2016.2610522.

Seethala, C., and Á. Horváth (2010), Global assessment of AMSR-E and MODIS cloud liquid water path retrievals in warm oceanic clouds, *J. Geophys. Res.*, 115, D13202, doi:10.1029/2009JD012662.

Seethala, C., Meirink, J. F., Horváth, Á., Bennartz, R., and Roebeling, R., 2018: Evaluating the diurnal cycle of South Atlantic stratocumulus clouds as observed by MSG SEVIRI, *Atmospheric Chemistry and Physics*, 18, 13 283–13 304, doi:10.5194/acp-18-13283-2018.

Stengel, M., Kniffka, A., Meirink, J. F., Lockhoff, M., Tan, J., and Hollmann, R., 2014: CLAAS: the CM SAF cloud property data set using SEVIRI, *Atmos. Chem. Phys.*, 14, 4297–4311.

Stephens, G. L., D. G. Vane, R. J. Boain, G. G. Mace, K. Sassen, Z. Wang, A. J. Illingworth, E. J. O'Connor, W. B. Rossow, S. L. Durden, S. D. Miller, R. T. Austin, A. Benedetti, C. Mitrescu, The CloudSat Science Team, 2002: The CloudSat mission and the A-Train, *Bull. Amer. Meteorol. Soc.*, 83, 12, 1771-1790.

van Diedenhoven, B., A. M. Fridlind, B. Cairns, A. S. Ackerman, and J. E. Yorks, 2016: Vertical variation of ice particle size in convective cloud tops, *Geophys. Res. Lett.*, 43, 4586–4593, doi:10.1002/2016GL068548.

Yang, P., L. Bi, B. A. Baum, K.-N. Liou, G. W. Kattawar, M. I. Mishchenko, and B. Cole, 2013: Spectrally consistent scattering, absorption, and polarization properties of atmospheric ice crystals at wavelengths from 0.2 to 100 μm , *J. Atmos. Sci.*, 70, 330-347, doi:10.1175/JAS-D-12-039.1.

13 Acronyms

AMSR2	Advanced Microwave Scanning Radiometer 2
AMSR-E	Advanced Microwave Scanning Radiometer for EOS
ATBD	Algorithm Theoretical Baseline Document
AVHRR	Advanced Very High Resolution Radiometer
BC-RMSD/ BC-RMSE	Bias-Corrected RMSD / RMSE
BT	Brightness Temperature
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarisation
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
CDO	Climate Data Operators
CDOP	Continuous Development and Operations Phase
CFC	Fractional Cloud Cover
IFOT	Integrated Cloud Optical Thickness
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
CRE	Cloud particle effective radius
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)
DARDAR	raDAR/iIDAR

DRR	Delivery Readiness Review
DWD	Deutscher Wetterdienst (German Met Service)
ECMWF	European Centre for Medium Range Forecast
ECV	Essential Climate Variable
ERA5	ECMWF Re-Analysis 5
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FAR	False Alarm Ratio
FCDR	Fundamental Climate Data Record
FCI	Flexible Combined Imager
GCOM-W	Global Change Observation Mission for Water
GCOS	Global Climate Observing System
GMI	Global Precipitation Measurement Microwave Imager
GSICS	Global Space-Based Inter-Calibration System
ICDR	Interim Climate Data Record
IR	InfraRed
ISCCP	International Satellite Cloud Climatology Project
ITCZ	Inter Tropical Convergence Zone
IWP	Ice Water Path
JAXA	Japan Aerospace Exploration Agency
JCH	Joint Cloud properties Histogram
KNMI	Koninklijk Nederlands Meteorologisch Instituut (Dutch Met Service)
KSS	Hanssen-Kuipers Skill Score
L1.5	Level 1.5
L2	Level 2
L3	Level 3
LUT	Look-Up Table
LWP	Liquid Water Path
MAC-LWP	Multisensor Advanced Climatology of Liquid Water Path (data record)

MODIS	Moderate Resolution Imaging Spectroradiometer
MSG	Meteosat Second Generation
MTG	Meteosat Third Generation
NIR	Near InfraRed
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
RMSE/ RMSD	Root Mean Square Error/Difference
RSS	Remote Sensing Systems
RTTOV	Radiative Transfer model for TOVS
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SAF	Satellite Application Facility
SMHI	Swedish Meteorological and Hydrological Institute
SSM/I	Special Sensor Microwave Imager
SYNOP	Synoptic observations
SZA	Solar Zenith Angle
TCDR	Thematic Climate Data Record
TMI	Tropical Rainfall Measurement Mission Microwave Imager
VIS	Visible
VZA	Viewing Zenith Angle