OPERATIONAL SUPPORT TO THE CROP MONITORING AND YIELD FORECASTING ACTIVITIES (MARSOP 6) Lot 2: Acquisition and Processing of Satellite Data

SC 7 Feasibility study – Quantile Mapping between FPAR products from VIIRS and MODIS

Deliverable D3 Final Report

07 February 2025

Clement Atzberger, Richard Kidd

VITO
Vlaamse Instelling voor Technologisch Onderzoek
Flemish Institute for Technological Research
Boeretang 200
B-2400 Mol, Belgium

EODC Earth Observation Data Centre for Water Resources Monitoring GmbH
Franz-Grill-Straße 9
1030 Wien
Austria

Version History

Version	Date	Prepared by	Reviewed by	Approved by	Signature
0.1	31/01/2025	Clement Atzberger			
0.2	07/02/2025	Richard Kidd			
0.3	07/02/2025	Clement Atzberger			
0.4	07/02/2025	Clement Atzberger	Isabelle Piccard		

TABLE OF CONTENT

List of Tables

List of Figures

List of Acronyms

BIAS	Mean difference between two samples
CV	Cross-validation
FAPAR	Fraction of Absorbed Photosynthetic Active Radiation
FTP	File Transfer Protocol
DELTA	Mean offset approach for inter-calibration
GDAL	Geospatial Data Abstraction Library
HDF	Hierarchical Data Format
LP DAAC	Land Processes Distributed Active Archive Center
LTA	Long Term Average
MAD	Mean absolute difference between two samples
MCYFS	MARS Crop Yield Forecasting System
MODIS	Moderate Resolution Imaging Spectroradiometer
MVC	Maximum Value Compositing
NRT	Near Real-Time
ORIG	Approach without any correction
QM	Quantile Mapping
QMF2	Quantile Mapping with ±2 dekads window size
PL	Production Line
POLY	Polynomial approach for inter-calibration
SDS	Scientific Data Sets
STD	Standard Deviation

Relevant Documents:

[Ref 1] OPERATIONAL SUPPORT TO THE CROP MONITORING AND YIELD FORECASTING ACTIVITIES (MARSOP 6) Lot 2: Acquisition and Processing of Satellite Data, 2nd Specific Contract, MODIS D1.1:Final report on results of tasks 1-4, 17 October 2023, SC2 (“Intercalibration of biophysical parameters”)

[Ref 2] OPERATIONAL SUPPORT TO THE CROP MONITORING AND YIELD FORECASTING ACTIVITIES (MARSOP 6) Lot 2: Acquisition and Processing of Satellite Data, 4th Specific Contract, MODIS D1: Final report, 12 January 2024, SC4 (“Preparation of a pre-operational production line”)

[Ref 3] MODIS VIIRS SC7 request (received 26-07-2024):

1 AIM OF SC7

The present study provides the necessary scientific basis to allow the JRC to take an informed decision about advantages and limitations of different FAPAR intercalibration strategies aligning the VIIRS FPAR time series (version 2) to the well-established MODIS (reference) time series.

The reported work is the result of a 4-month feasibility study within MARSOP6, SC7. It builds on two earlier feasibility studies, namely SC2 ("Intercalibration of biophysical parameters")[Ref 1] and SC4 ("Preparation of a pre-operational production line")[Ref 2].

The work reported here was split into four tasks, as specified in the MODIS VIIRS SC7 request (received 26-07-2024)[Ref 3]:

- Task 1: Implement quantile mapping
- Task 2: Compare and discuss quantile mapping against mean difference correction
- Task 3: Highlight opportunities and limitations of alternative procedures
- Task 5: Test a novel quantile mapping approach based on 2D polynomials

The current study shall in particular:

1. evaluate the quantile mapping (QM) intercalibration results on a set of sample points identical to SC4, but using the new, recently released, version V002 of VIIRS products. In particular, the study shall complement the period 2018-2023 of VNP+VJ1 with the period 2012-2017 of VNP (all version 2), so that the quantile mapping can be calibrated on the full length of the time series (2012-2023) in order to enable a better intercalibration between the VNP+VJ1 smoothed outputs as compared to the current one made using the mean difference. As QMF2 (dekadal quantile mapping with window size 2, 5 dekads in total, method equant, qstep 0.01) was found to be the best intercalibration method on a previous study on the intercalibration of the full time series (from 2012) of VNP v1, this hyperparameter setting should be applied;
2. compare quantile mapping (QM) results in the window 2018-2023 over identical sample points, as it cannot be given for granted that QM trained on the full time series (2012-2023) outperforms the mean difference method trained on the shorter period (2018-2023). This will permit to evaluate in a scientifically sound way the opportunity of using QM;
3. compare the QM results obtained with the proposed hyperparameter setting outline above (QMF2) against a computationally more efficient procedure where FAPAR differences between sensors are directly modeled as a 2-dimensional polynom of dekad (first predictor variable) and VIIRS FAPAR (second predictor variable).

2 DATA AND METHODS

To provide robust results, the study was conducted on a sparsely sampled global grid with size 699 x 1920 pixel (**Figure 2-1** - top). This global grid was derived from the full resolution (500m) FAPAR dataset by sub-sampling every 42nd sample.

The VIIRS and MODIS FAPAR datasets were produced by EODC and cover the time period from Dekad 19 in 2012 till Dekad 18 in 2023. This yields for each of the two sensors (VIIRS and MODIS) a FAPAR time series of length 11 years x 36 dekads. Per dekad, year and consolidation stage (C0 to CF) 324,004 (vegetated) pixels have been considered. FAPAR values rank between 0 and 100.

For completeness and comparability with previous reports, from this global dataset, 70 locations within Europe and Africa were extracted and reported separately (**Figure 2-1**- bottom). These results are presented in the section 6; Annex: Supplementary Results.

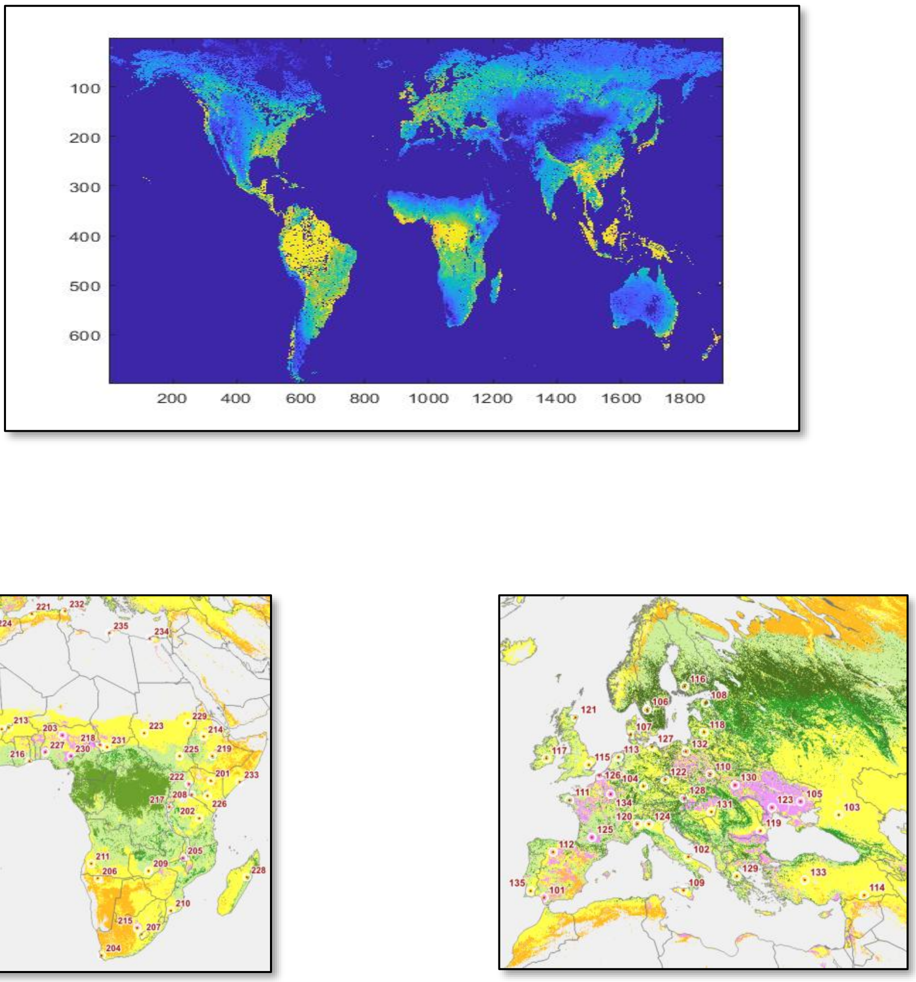


***Figure 2-1**. (top) global grid analyzed for the purpose of this FAPAR intercalibration study. The global window has the size 699 x 1920 pixels. Background pixels (e.g., ocean and bare surfaces) are shown in dark blue. (bottom) specific locations analyzed within Africa, respectively, within Europe. Results referring to these few 70 samples are reported separately in the Annex.*

Building on previous studies (SC2 and SC4, [Ref 1],[Ref 2]), four FAPAR intercalibration approaches were assessed:

- No correction (ORIG)
- Offset correction (DELTA)
- Quantile mapping (QMF2)
- 2-dimensional quantile mapping (POLY)

To avoid problems related to overfitting, and to guarantee that the results reflect well the expected accuracy of the implemented method (if any), a 6-fold leave one out approach was implemented across all intercalibration methods. This implies that from the 11-year FAPAR dataset, always one year between 2018 and 2023 was left-out. The principle is illustrated in **Figure 2-2**.

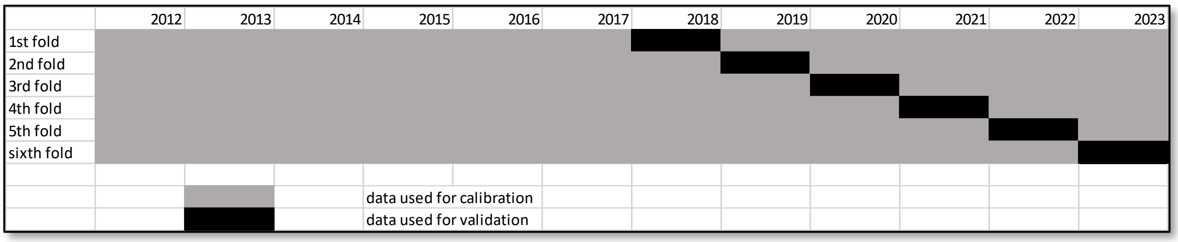


Figure 2-2 *Illustration of the 6-fold leave one out approach implemented for analyzing FAPAR time series from VIIRS and MODIS. In light gray are shown the years included in the training dataset. Data used for validation are shown in black. In the six folds, each of the six validation years (2018 to 2023) is left out once. The cross-validation obviates overfitting problems. Only those cross-validated results are reported in this study (2018-2023 with a total of 216 dekads).*

The implemented 6-fold validation yields cross-validated results. Only these results will be reported here, implying that the final validation dataset has (per consolidation stage) a size of 36 dekads x 6 years x 324,004 pixels. The cross-validated statistics are reflecting the accuracy one can expect in forthcoming years when implementing the discussed FAPAR intercalibration approaches. This is a major change with respect to SC2 and SC4, which were solely reporting calibration results.

2.1 Offset correction - DELTA

In the offset correction (labeled as 'DELTA'), one establishes per dekad the average difference between the (original) VIIRS FAPAR data and the corresponding (reference) MODIS data using the calibration data from the respective fold:

$$\Delta_{dekad} = MODIS_{dekad}^{cal} - VIIRS_{dekad}^{cal} \qquad (Equation\ 1)$$

This correction (Δ) is applied on the left-out data ($VIIRS^{val}$) to generate a corrected VIIRS:

$$VIIRS_{dekad}^{corrected} = VIIRS_{dekad}^{val} + \Delta_{dekad} \qquad (Equation\ 2)$$

The calculations are done pixel-by-pixel and separately for each of the six consolidation stages. Per fold, one Δ is calculated. Implementing this approach implies storing one offset value (Δ) per pixel, consolidation stage and dekad.

2.2 Quantile mapping - QMF2

In the quantile mapping (labeled as 'QMF2'), one establishes for each dekad a transfer function between the sorted VIIRS FAPAR data and the corresponding (sorted) MODIS data (**Figure 2-3**). To minimize interpolation needs, a total of 101 (regularly spaced) quantiles are established. To increase the number of available data points, the two preceding and two following dekads are included in the analysis for each of the 36 dekads.

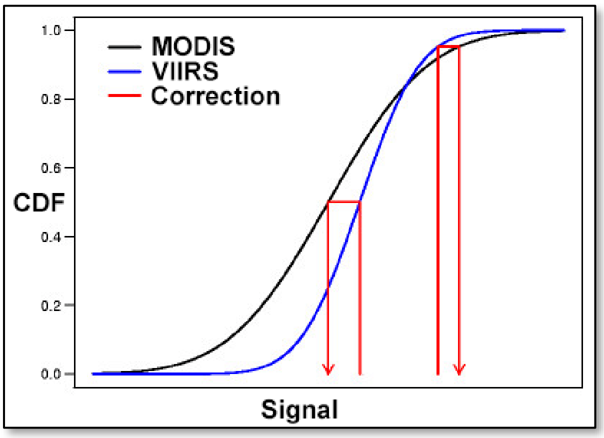


Figure 2-3 *Illustration of the quantile mapping (QM). Both FAPAR signals (e.g. from VIIRS and MODIS) are sorted so that the matching between VIIRS and MODIS can be established (indicated by the red arrows).*

For each of the VIIRS FAPAR values, one MODIS value is derived having the same CDF value (cumulated distribution function). This provides the necessary matching between the (uncorrected) VIIRS FAPAR and the reference MODIS value.

The calculations are done pixel-by-pixel and separately for each of the six consolidation stages. Per fold, a separate quantile mapping is established. Implementing this approach implies that per pixel, consolidation stage and dekad, a table with 101 x 2 entries needs to be stored: one column for the (sorted) VIIRS FAPAR data (input), one for the output (e.g., the corrected VIIRS FAPAR data).

2.3 2D Polynomial fitting - POLY

Instead of working dekad by dekad independently, it is possible to fit a 2D polynomial so that the corrected VIIRS FAPAR is a function of the uncorrected VIIRS data ($VIIRS_{cal}$) for all 36 dekads simultaneously (labeled as 'POLY'):

$$VIIRS_{1:36}^{corrected} = VIIRS_{1:36}^{val} + \Delta_{1:36} \quad \text{(Equation 3)}$$

An illustration of the 2-dimensional offset (Δ) is shown in Figure 2-4 for a polynomial of degree 2 and 3; e.g. quadratic over time (dekad) and cubic in VIIRS FAPAR. Such a polynomial has nine coefficients (p_{00} to p_{03}) and can be written as follows (X: dekad; Y: VIIRS FAPAR):

$$\Delta = p_{00} + p_{10}X + p_{01}Y + p_{20}X^2 + p_{11}XY + p_{02}Y^2 + p_{21}X^2Y + p_{12}XY^2 + p_{03}Y^3 \quad \text{(Equation 4)}$$

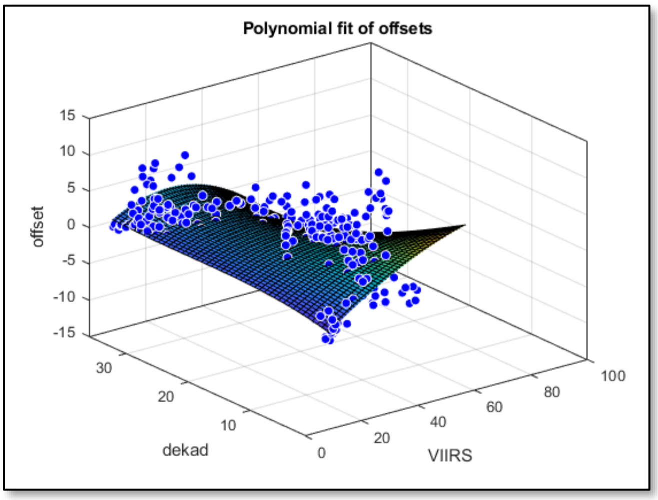


Figure 2-4. *Illustration of a 2-dimensional polynomial mapping. In the example, a quadratic/cubic polynomial is fitted (poly$_{23}$) - see Equation 4. To establish the FAPAR offset, both FAPAR signals (e.g. from VIIRS and MODIS) are sorted within each dekad and the difference between the two sorted signals is modeled. In the graph, the dots represent the observations and the curved surface the fitted 2-dimensional polynomial*

To ensure that the derived surface is continuous between dekad 36 (end of December) and dekad 1 (start of January), when fitting the polynomial, two dekads are padded to the start and end of the time series, respectively. The resulting sequence is thus: [35-36-01-02-03-…. -34-35-36-01-02].

The calculations are done pixel-by-pixel and separately for each of the six consolidation stages. Per fold, a single 2D polynomial is established across all dekads.

Implementing this approach implies that per pixel and consolidation stage, only nine coefficients need to be estimated and stored (p_{00} to p_{03} in the case of a poly$_{23}$). To derive the offset specific for a given dekad, the dekad number (X) needs to be entered together with the (uncorrected) VIIRS FAPAR value (Y) in Equation 4.

2.4 No correction - ORIG

To assess how well VIIRS and MODIS match without any correction, the two FAPAR time series were also directly compared without any further adjustments. To ensure perfect comparability, the same time period was analyzed (e.g., 2018-2023). This provides a suitable baseline and upper bound against which the three intercalibration methods can be compared.

3 RESULTS

The cross-validated mean absolute differences (MAD_{cv}) across the six validation years are exemplarily shown in **Figure 3-1** for the least consolidated FAPAR product (C0). The polynomial approach (POLY) performs best, followed by the offset correction (DELTA) and the quantile mapping (QMF2). All three intercalibration methods significantly improve the matching between VIIRS and MODIS FAPAR compared to the original FAPAR data (ORIG).

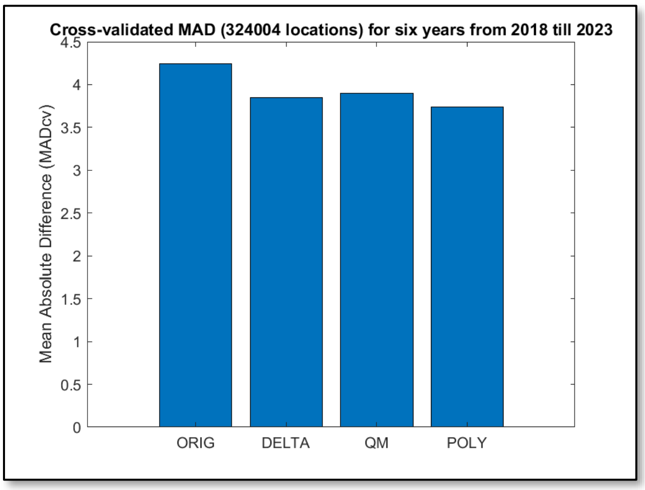


Figure 3-1. *Cross-validated mean absolute differences (MADcv) for the six validation years 2018-2023 and the three analyzed intercalibration methods (DELTA to POLY). For comparison, the baseline method (ORIG) is also added. The analysis is based on FAPAR data from the C0 stage and involves FAPAR data from 324.004 pixels, 36 dekads and 6 years*

For the uncorrected data (ORIG) as well as the three intercalibration methods, the mismatch between VIIRS and MODIS FAPAR time series consistently decreases with increasing consolidation stage (e.g., from C0 to CF) (**Figure 3-2** and **Table 3-1**). The relative ranking between the three intercalibration methods, however, remains largely unchanged. In all but one case (e.g. for CF), POLY performs better than DELTA and, in all cases, better than QMF2. The systematic bias ($BIAS_{cv}$) is generally very low (< 1 FAPAR unit) even for the uncorrected data (**Table 3-1**).

Table 3-1. *Cross-validated (2018-2023) mean absolute differences (MAD_{cv}) (left) and mean differences ($BIAS_{cv}$) (right) between MODIS and VIIRS FAPAR for different intercalibration techniques and the six consolidation stages (C0 to CF). In bold, the best results*

	MAD_{cv}					$BIAS_{cv}$			
	ORIG	**DELTA**	**QMF2**	**POLY**		**ORIG**	**DELTA**	**QMF2**	**POLY**
C0	4,24	3,84	3,90	**3,74**		0,50	**0,32**	0,68	0,56
C1	3,66	3,19	3,26	**3,11**		0,55	**0,32**	0,68	0,48
C2	3,36	2,84	2,93	**2,78**		0,60	**0,31**	0,69	0,41
C3	3,30	2,73	2,85	**2,69**		0,65	**0,29**	0,69	0,35
C4	3,32	2,72	2,86	**2,70**		0,68	**0,27**	0,69	0,30
CF	3,34	**2,58**	2,80	2,62		0,80	0,16	0,71	**0,01**
Average	3,54	2,99	3,10	**2,94**		0.63	**0.28**	0.69	0.35
$\Delta_{(max-min)}$	**0.90**	1,26	1,10	1,12		0.30	0.17	**0.03**	0.55

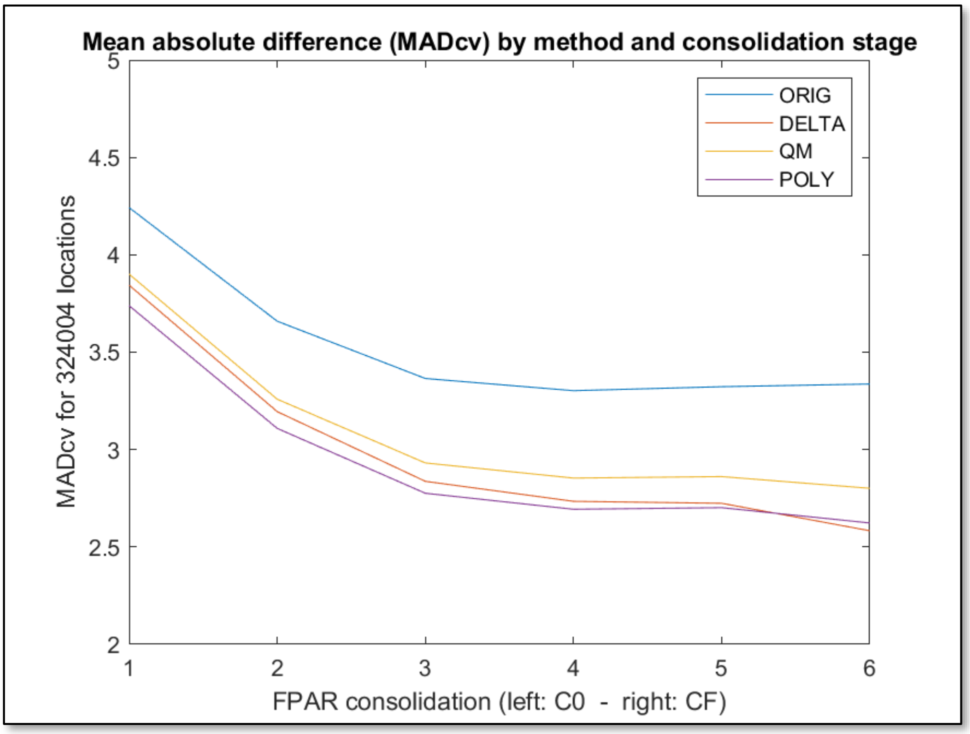


Figure 3-2. *Impact of the consolidation stage (from C0 left to CF on the right) on the cross-validated mean absolute differences (MADcv) for the six validation years 2018-2023. The analysis is based on FAPAR data from 324.004 pixels, 36 dekads and 6 years. The three intercalibration methods are DELTA, QMF2 and POLY. The uncorrected baseline is ORIG*

For two exemplary consolidation stages (e.g., C0 and C4), the pie-charts in **Figure 3-3** show the relative frequency of the best performing method across all pixel, confirming the above statements; for the C0 stage (left), POLY leads to the lowest cross-validated RMSE in 50% of the cases when analyzing data from six validation years, 36 dekads and 324,004 pixels. In another one third of the cases, DELTA is the best method to align VIIRS and MODIS FAPAR time series.

This trend shifts when analyzing the C4 stage (right). Here, DELTA is slightly more often selected as the best performing method (43%) compared to POLY (38%). For all consolidation stages, the use of quantile mapping (QMF2) was only fruitful in about 6-8% of the cases. The results for the other consolidation stages are found in the lower part of **Figure 3-3**, confirming the mentioned trends.

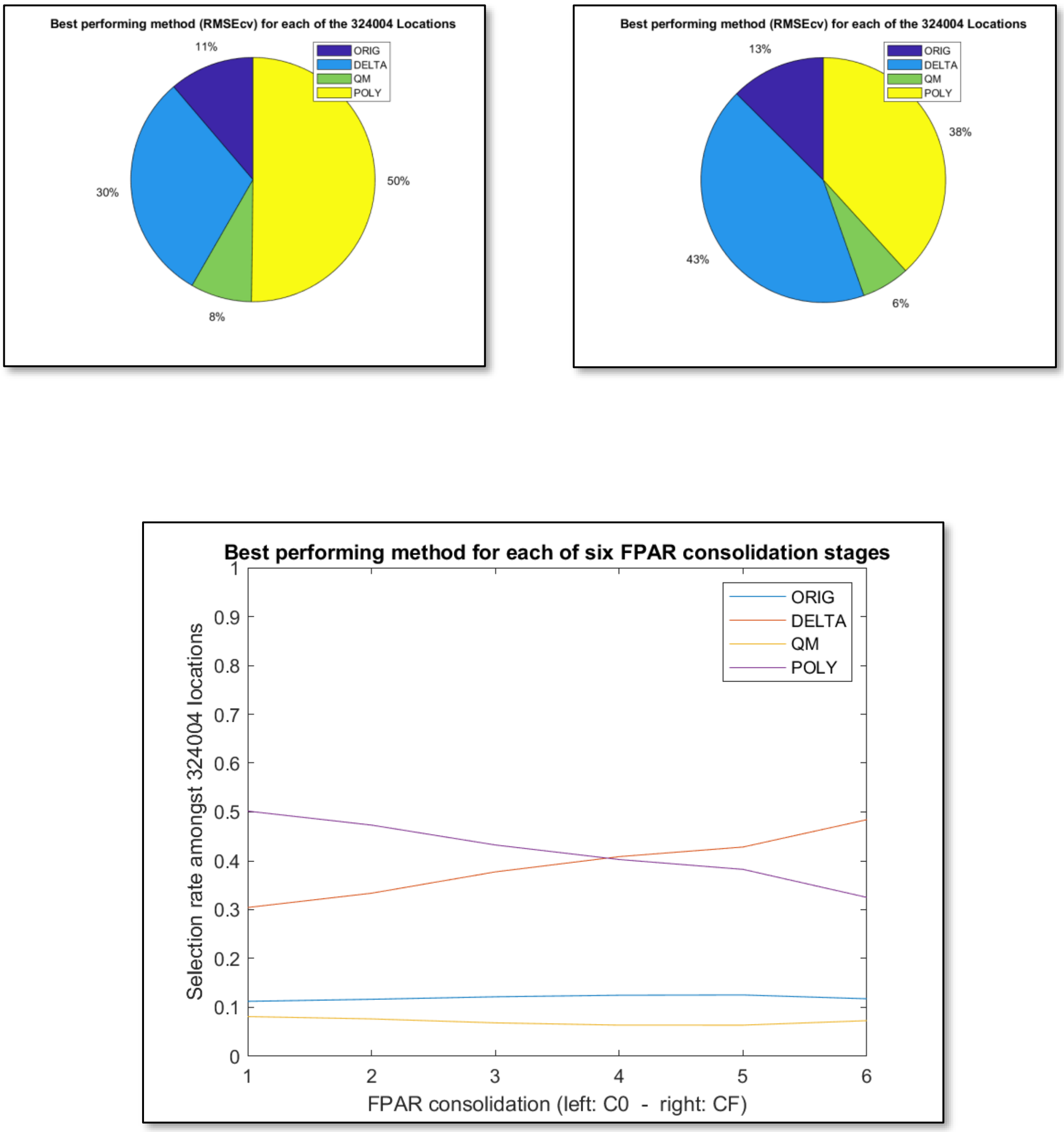


Figure 3-3. *Relative frequency with which each method leads to the lowest mismatch between VIIRS and MODIS FAPAR based on the cross-validated RMSE. (top) pie-charts for FAPAR consolidation stage C0 (left) and CF (right). (bottom) Selection rates for all 6 FAPAR consolidation stages and methods (from C0 left to CF right). The results are based on the analysis of 324004 pixel, 36 dekads and 6 validation years*

For each of the six validation years 2018 to 2023, **Figure 3-4** shows the cross-validated mean absolute difference (MAD_{cv}) of POLY and DELTA compared to the original FAPAR data (ORIG). For the first four consolidation stages (C0 to C3) (first two rows), the relative ranking between POLY and DELTA remains unchanged over the six years, with consistent – but however steadily decreasing – advantages of POLY.

Only for C4 and CF (last row), both alternatives lead to more or less similar matchings. However, for these two consolidation stages (e.g., C4 and CF), even the uncorrected FAPAR data (ORIG) is already performing better (e.g., with a MAD_{cv} of around 3.3) compared to the overall best results obtained for C0 (e.g., a MAD_{cv} of around 3.7 for POLY).

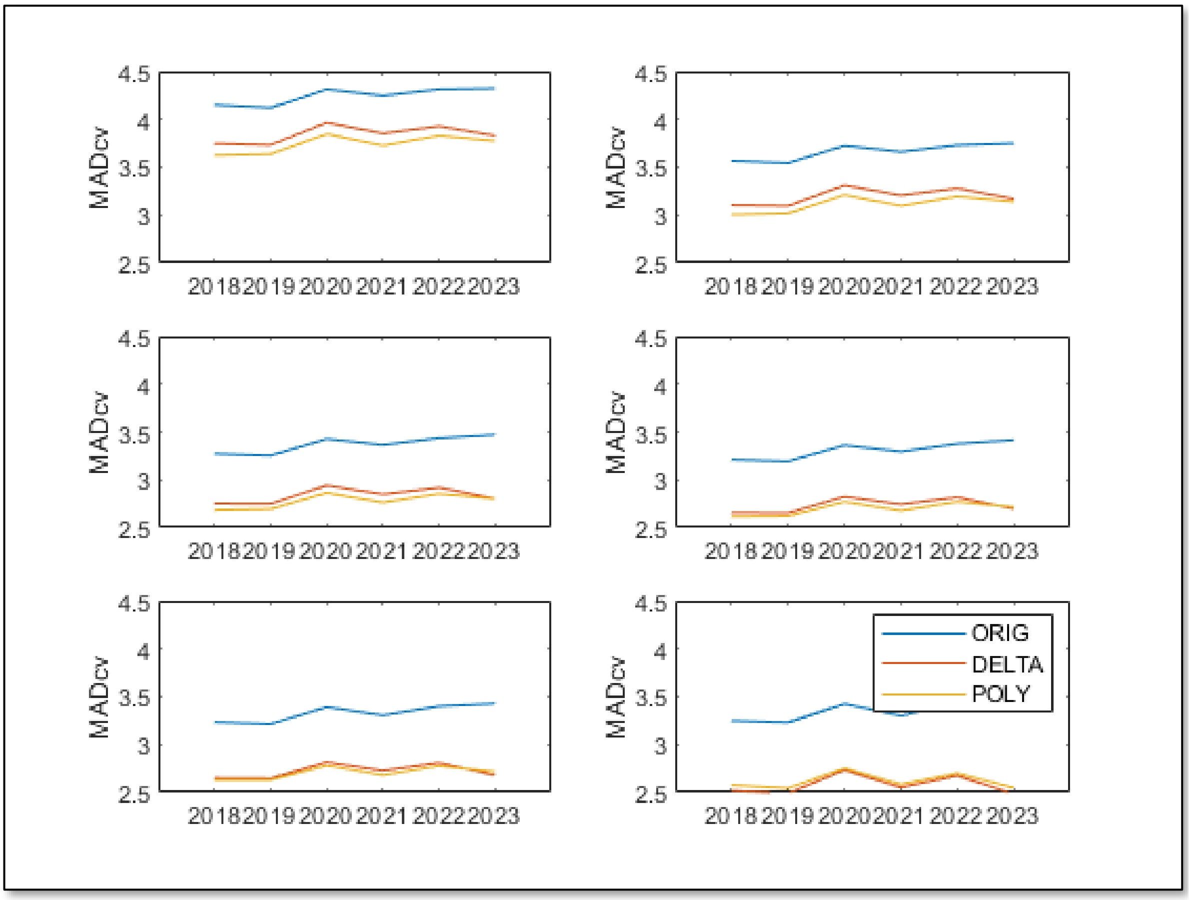


Figure 3-4. *Cross-validated MAD between VIIRS and MODIS separately for each of the six validation years (2018 to 2023) and the two intercalibration methods DELTA and POLY, compared to the uncorrected data (ORIG). Results are ordered from C0 top left, C1 top right,…, to CF on the bottom right. Each analysis involves the analysis of 324,004 pixels and 36 dekads*

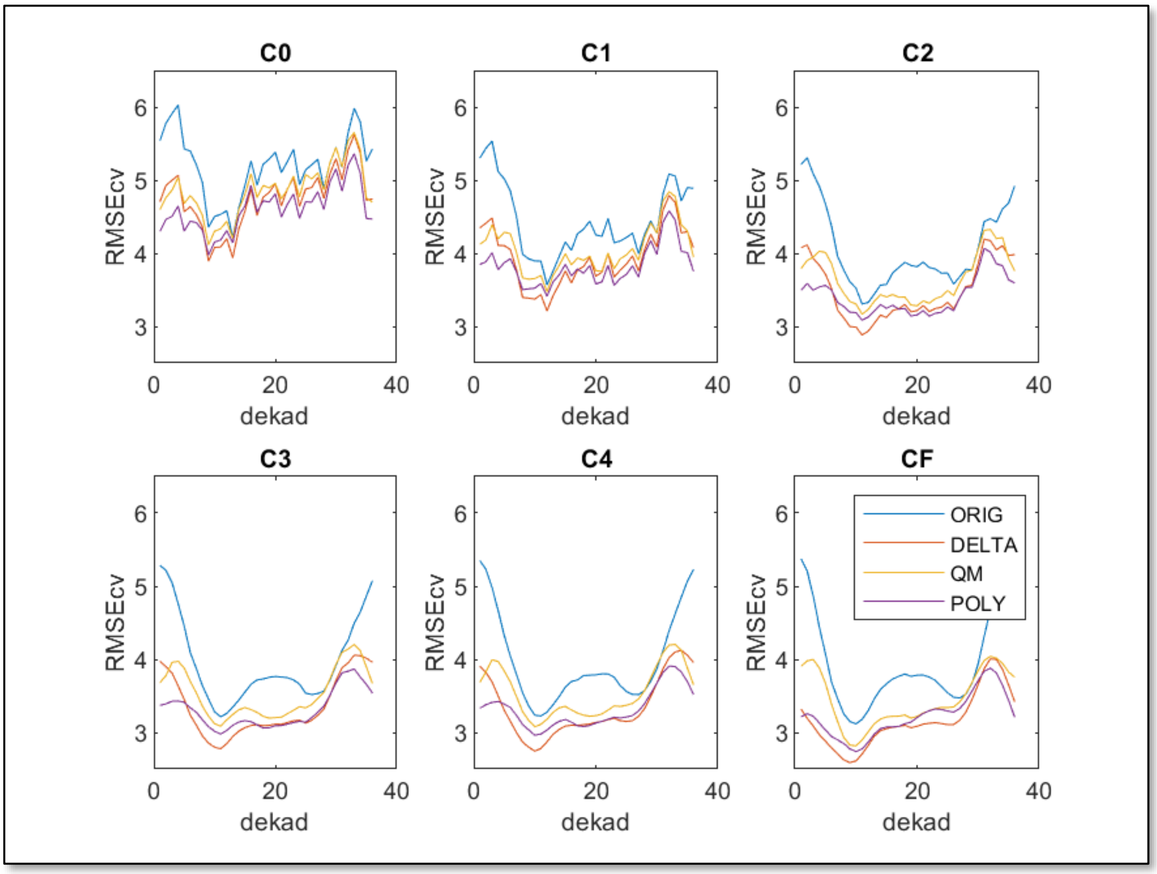


Figure 3-5. *Cross-validated RMSE between VIIRS and MODIS FAPAR by dekad for the six consolidation stages (C0 to CF) and the six validation years 2018-2023. Per dekad and consolidation stage, the analysis is based on FAPAR data from 324,004 pixels and 6 years. The three FAPAR intercalibration methods are DELTA, QMF2 and POLY. The uncorrected baseline is ORIG*

The above described trends largely remain unchanged if the analysis is done for each of the 36 dekads independently (**Figure 3-5**). With the exception of CF, POLY generally outperforms DELTA (and the quantile mapping). For all methods and most consolidation stages, the highest mismatch in the two FAPAR time series occurs around December and the lowest errors around April (**Table 3-2**).

The advantages of POLY are most pronounced when the two FAPAR time series are the least aligned. For the FAPAR consolidation stages C0 to C2, POLY outperforms DELTA for all months from July to February (**Table 3-2**). On the contrary, DELTA has slight advantages during March-May, where the mismatch even for the uncorrected FAPAR (ORIG) is relatively low.

In the cases where the two FAPAR series are not well aligned, it seems that the distributions are not only shifted (e.g., could be corrected by a simple offset – DELTA) but that at the same time, the spread of the distributions is affected. This latter effect cannot be corrected by DELTA but must be handled using appropriate methods such as QMF2 or POLY. This also explains why QMF2 usually underperforms DELTA, except for the months with the highest mismatch between the VIIRS and MODIS FAPAR time series (see for example JAN for the CO product; **Table 3-2**). As QMF2 suffers from overfitting due to its extreme flexibility, POLY always and consistently outperforms QMF2.

Table 3-2. *Cross-validated mean absolute difference (MADcv) between VIIRS and MODIS FAPAR by month for the six consolidation stages (C0 to CF) and the four methods. In bold, the best results for each of the twelve months.*

						C0						
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
ORIG	4,74	4,65	3,94	3,69	3,72	4,18	4,29	4,24	4,23	4,24	4,56	4,52
DELTA	3,96	3,86	3,43	**3,34**	**3,47**	**3,83**	3,91	3,94	4,02	4,14	4,36	4,05
QMF2	3,79	3,89	3,55	3,48	3,64	3,97	3,93	3,95	4,09	4,23	4,43	4,02
POLY	**3,57**	**3,60**	**3,43**	3,43	3,62	3,85	**3,79**	**3,78**	**3,86**	**4,04**	**4,20**	**3,83**
						C1						
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
ORIG	4,52	4,17	3,40	3,10	3,23	3,47	3,55	3,50	3,41	3,55	4,07	4,01
DELTA	3,60	3,32	**2,85**	**2,72**	**2,92**	**3,05**	3,09	3,11	3,16	3,40	3,83	3,47
QMF2	3,39	3,43	3,02	2,89	3,09	3,15	3,09	3,15	3,23	3,51	3,88	3,42
POLY	**3,16**	**3,12**	2,92	2,87	3,05	3,06	**2,99**	**3,00**	**3,06**	**3,34**	**3,67**	**3,22**
						C2						
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
ORIG	4,39	3,91	3,14	2,79	2,91	3,16	3,20	3,13	3,03	3,19	3,65	3,96
DELTA	3,31	3,00	**2,54**	**2,40**	**2,55**	2,67	2,65	2,65	2,74	3,02	3,39	3,32
QMF2	3,12	3,22	2,78	2,60	2,75	2,76	2,67	2,72	2,86	3,15	3,49	3,22
POLY	**2,85**	**2,86**	2,67	2,58	2,68	**2,67**	**2,59**	**2,59**	**2,71**	**2,99**	**3,26**	**3,03**
						C3						
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
ORIG	4,40	3,75	3,05	2,69	2,85	3,09	3,14	3,08	2,93	3,08	3,55	4,10
DELTA	3,19	2,81	**2,43**	**2,30**	**2,47**	2,54	2,55	2,59	**2,63**	2,91	3,27	3,32
QMF2	3,06	3,13	2,71	2,52	2,68	2,65	2,59	2,67	2,78	3,06	3,39	3,16
POLY	**2,75**	**2,75**	2,60	2,47	2,58	**2,54**	**2,52**	**2,56**	2,65	**2,91**	**3,16**	**3,00**
						C4						
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
ORIG	4,42	3,69	2,98	2,68	2,89	3,11	3,17	3,10	2,93	3,07	3,62	4,28
DELTA	3,11	**2,71**	**2,37**	**2,30**	**2,49**	2,55	2,57	2,62	**2,64**	**2,90**	3,31	3,35
QMF2	3,10	3,12	2,66	2,52	2,70	2,65	2,61	2,71	2,79	3,07	3,40	3,15
POLY	**2,73**	2,74	2,55	2,46	2,58	**2,54**	**2,55**	**2,61**	2,68	2,92	**3,19**	**3,02**
						CF						
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
ORIG	4,52	3,48	2,74	2,63	2,97	3,15	3,18	3,10	2,90	3,06	3,89	4,52
DELTA	**2,61**	**2,35**	**2,16**	**2,21**	**2,48**	2,54	**2,56**	**2,59**	**2,59**	**2,87**	3,26	2,99
QMF2	3,28	2,99	2,38	2,33	2,57	2,61	2,64	2,72	2,76	3,02	3,27	3,17
POLY	2,63	2,46	2,32	2,28	2,49	**2,53**	2,64	2,73	2,73	2,95	**3,14**	**2,80**

This trend is further exemplified in **Figure 3-6**. The figure shows how often (in relative terms) POLY outperforms DELTA as a function of the original mis-alignment between (uncorrected) VIIRS and MODIS FAPAR. This deviation between the two time series is shown on the x-axis, whereas the fraction of POLY “winning” over DELTA is shown on the y-axis.

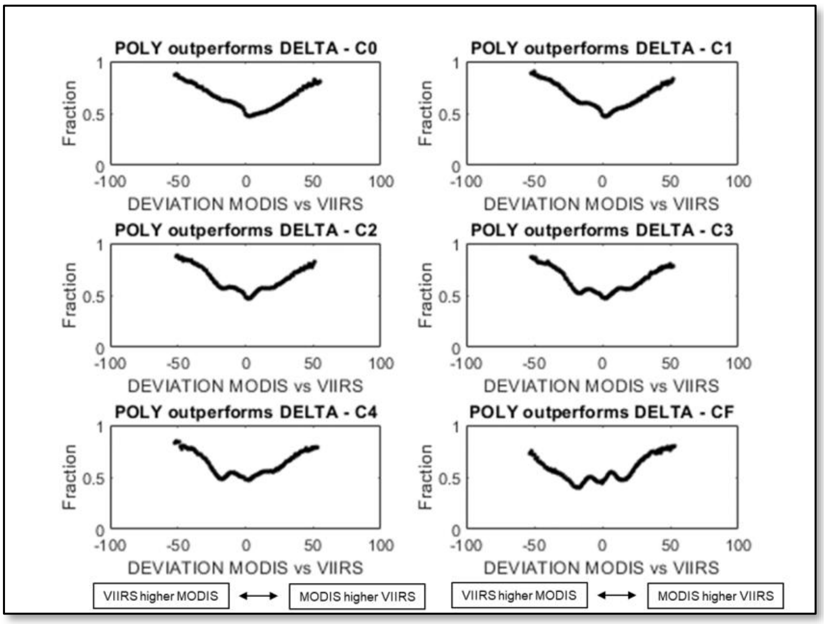


Figure 3-6. *Performance comparison between POLY and DELTA methods. For each consolidation stage, the fraction is shown in which POLY yields smaller MADcv compared to DELTA. This fraction (y-axis) is shown as a function of the (uncorrected) mis-alignment between MODIS and VIIRS FAPAR time series (x-axis).*

Interestingly, with small mis-alignments of VIIRS and MODIS FAPAR time series (e.g., deviations at around ±10 FAPAR), POLY and DELTA perform almost equally well (e.g., the fraction of POLY being selected is close to 50%) – and this trend holds for all FAPAR consolidation stages. However, with larger deviations - and in particular when the (uncorrected) VIIRS FAPAR is much higher than the MODIS FAPAR - POLY excels more and more over DELTA, reaching fractions in excess of 90%.

For time series analysis – and in particular for sensor intercalibration studies – it is not only important that the mismatch between the two time series is low/minimal, but also that the mismatch is stable over the course of a year. This seasonal stability is shown in **Table 3-3**. The table reports for each of the four methods the difference between the highest and lowest dekadal MAD_{cv}. Lower values indicate a more stable behavior; e.g. no “seasonality”. In this respect, POLY shows clear advantages compared to the other approaches, and this across all six FAPAR consolidation stages. For most consolidation stages, QMF2 has a lower seasonal amplitude in errors compared to DELTA. ORIG shows the worst performance.

Table 3-3*. Intra-annual maximum difference in the dekadal cross-validated MAD between VIIRS and MODIS separately for each consolidation stage (MADmax - MADmin). Each analysis involves the analysis of 324,004 pixels, 36 dekads and six years (2018-2023). In* ***bold****, the best results.*

	C0	C1	C2	C3	C4	CF
ORIG	0,34	0,43	0,49	0,53	0,56	0,64
DELTA	0,30	0,32	0,34	0,35	0,35	0,31
QMF2	0,27	0,27	0,28	0,28	0,28	0,33
POLY	**0,23**	**0,21**	**0,21**	**0,21**	**0,22**	**0,25**

Comparing the FAPAR intercalibrations from POLY and DELTA on the (sparse) global grid, significant patterns emerge (**Figure 3-7**). The maps for each of the six consolidation stages show in yellow the locations where POLY outperforms DELTA and in blue where DELTA performs best. Clear spatial patterns can be seen reflecting the interplay between climate and dominant land cover; both obviously being intimately intertwined.

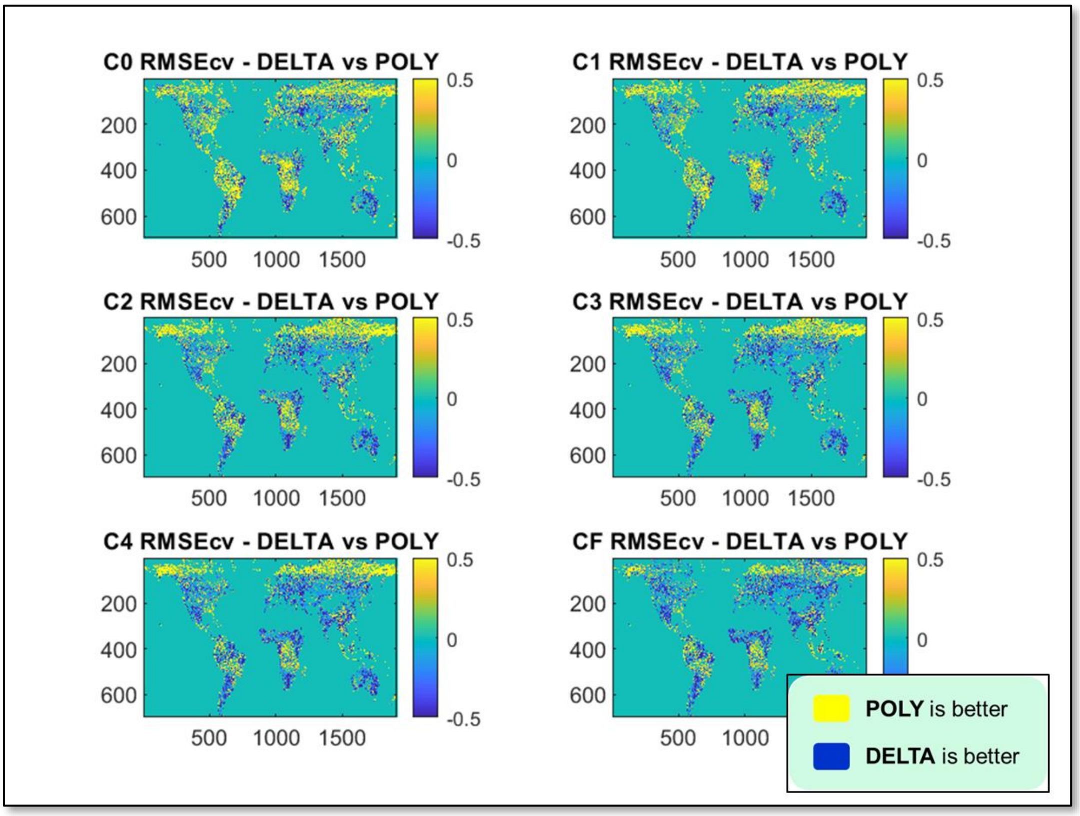


***Figure 3-7.** Maps indicating for each of the six consolidation stages (C0 to CF) if POLY outperforms DELTA in aligning the VIIRS and MODIS FAPAR time series (in yellow), or vice versa (in blue). For each of the maps, FAPAR data from six years and 36 dekads are analyzed*

In general terms, POLY usually outperforms DELTA in regions with less favorable observation conditions such as the boreal regions and the tropics (see also **Figure 3-6**). As discussed previously, this advantage vanishes when better consolidated FAPAR time series are analyzed (e.g. C4 or CF).

To disentangle the impact of the dominant land cover, the analysis was repeated for the eight major land cover classes shown in **Figure 3-8**.

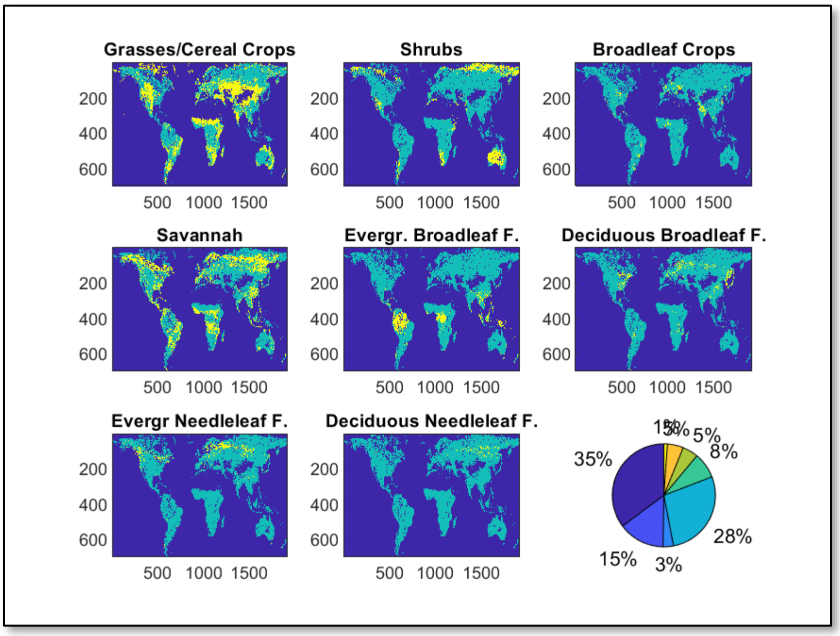


Figure 3-8. *Maps showing (in yellow) the respective dominant land cover class on the sparse global grid analyzed for this study. The relative proportions are shown in the pie chart starting at 12:00 in counter-clockwise order*

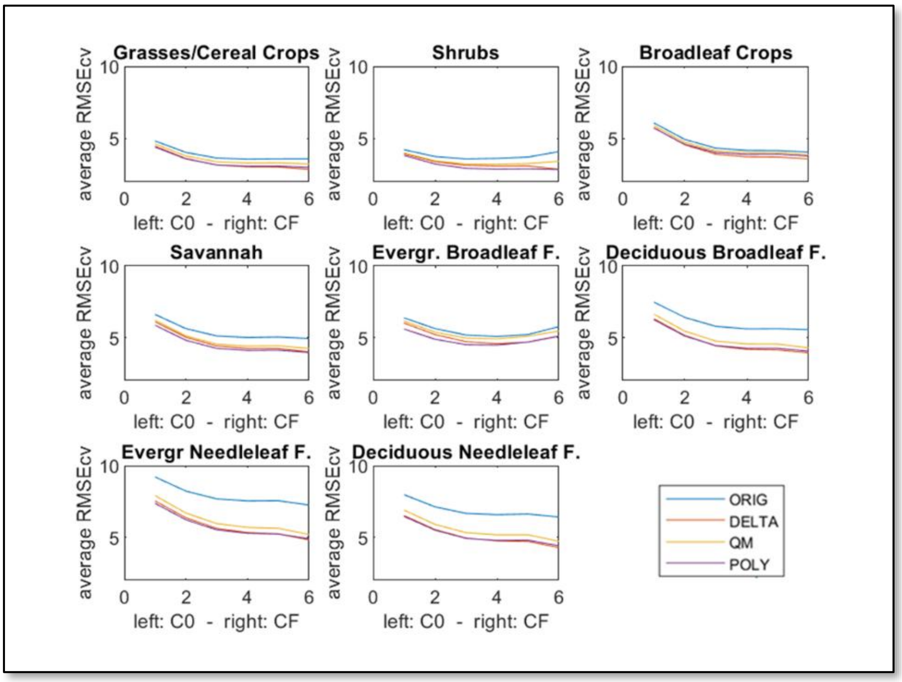


Figure 3-9. *Cross-validated RMSE for eight major land cover classes and different FAPAR intercalibration methods. In each sub-graph, the FAPAR consolidation stages are shown from left (C0) to right (CF). For the spatial distribution and relative frequency of the different land cover classes, see Figure 3-8*

Table 3-4. *Cross-validated mean absolute differences (MADcv) between VIIRS and MODIS FAPAR time series by land cover type and for the six FAPAR consolidation stages (C0 to CF): GCC – Grasses & Cereal Crops; SHR – Shrubs; BLC – Broadleaf Crops; SAV – Savannah; EBF – Evergreen Broadleaf Forest; DBF – Deciduous Broadleaf Forest; ENF – Evergreen Needleleaf Forest; DNF – Deciduous Needleleaf Forest. The number of locations for each of the eight land cover classes are also indicated (Nobs). In bold, for each land cover class the best performing intercalibration method*

	GCC	SHR	BLC	SAV	EBF	DBF	ENF	DNF
N_{obs}	113,995	47,125	10,634	90,283	26,112	16,034	16,248	3,573
				C0				
ORIG	3,23	3,05	4,26	4,65	3,15	5,29	6,84	6,00
DELTA	2,99	2,89	**3,98**	4,23	3,21	**4,45**	5,42	4,73
QMF2	3,10	2,92	4,10	4,24	3,28	4,63	5,56	4,93
POLY	**2,97**	**2,81**	4,00	**4,06**	**3,03**	4,47	**5,31**	**4,69**
				C1				
ORIG	2,71	2,78	3,45	3,94	2,65	4,44	6,01	5,32
DELTA	**2,41**	2,53	**3,14**	3,48	2,68	**3,67**	4,54	4,02
QMF2	2,56	2,57	3,29	3,50	2,78	3,83	4,68	4,21
POLY	2,43	**2,39**	3,22	**3,34**	**2,54**	3,67	**4,46**	**3,99**
				C2				
ORIG	2,46	2,67	3,06	3,58	2,33	3,95	5,56	4,90
DELTA	**2,11**	2,32	**2,71**	3,06	2,33	**3,15**	4,00	3,57
QMF2	2,30	2,42	2,90	3,11	2,47	3,35	4,15	3,78
POLY	2,16	**2,17**	2,83	**2,94**	**2,25**	3,19	**3,94**	**3,55**
				C3				
ORIG	2,41	2,71	2,96	3,50	2,19	3,80	5,42	4,83
DELTA	**2,03**	2,27	**2,61**	2,94	2,18	**2,98**	3,79	**3,40**
QMF2	2,25	2,42	2,82	3,02	2,37	3,20	3,95	3,64
POLY	2,11	**2,13**	2,76	**2,84**	**2,15**	3,05	**3,75**	3,42
				C4				
ORIG	2,42	2,78	2,96	3,51	2,19	3,81	5,42	4,87
DELTA	**2,03**	2,26	**2,61**	2,92	2,19	**2,95**	3,71	**3,37**
QMF2	2,26	2,46	2,83	3,01	2,39	3,17	3,87	3,63
POLY	2,12	**2,15**	2,77	**2,83**	**2,19**	3,04	**3,71**	3,42
				CF				
ORIG	2,34	2,93	2,89	3,40	2,43	3,77	5,28	4,79
DELTA	**1,89**	**1,96**	**2,52**	2,68	2,42	**2,80**	**3,41**	**3,11**
QMF2	2,17	2,50	2,73	2,85	2,59	2,98	3,58	3,36
POLY	2,04	2,06	2,67	**2,65**	**2,42**	2,89	3,46	3,21

In general terms, and across all FAPAR intercalibration techniques and FAPAR consolidation stages, FAPAR time series over shrubs are best aligned, while evergreen and deciduous needleleaf forests show the highest cross-validated RMSE (**Figure 3-9**). For most of the consolidation stages and land cover types, POLY performs best, closely followed by DELTA (**Table 3-4**).

To assess if potentially even better fitting 2-dimensional polynomials can be found, an ablation study was performed in which - for the C2 consolidation stage - four alternative polynomials were evaluated on the same (global) dataset. To ensure a perfect comparability, the exact same process as described before was used for data analysis.

Besides poly23 with nine coefficients (**Equation 5a**), four alternative polynomials with six to twelve coefficients were studied (X: dekad; Y: VIIRS FAPAR) (**Equation 5b** to **5e**):

poly23
(nine coefficients)

$$sf(x,y) = p_{00} + p_{10}\,X + p_{01}\,Y + p_{20}\,X^2 + p_{11}\,X\,Y + p_{02}\,Y^2 + p_{21}\,X^2\,Y + p_{12}\,X\,Y^2 + p_{03}\,Y^3 \quad \text{(Equation 5a)}$$

poly22
(six coefficients)

$$sf(x,y) = p_{00} + p_{10}\,X + p_{01}\,Y + p_{20}\,X^2 + p_{11}\,X\,Y + p_{02}\,Y^2 \quad \text{(Equation 5b)}$$

poly24
(twelve coefficients)

$$sf(x,y) = p_{00} + p_{10}\,X + p_{01}\,Y + p_{20}\,X^2 + p_{11}\,X\,Y + p_{02}\,Y^2 + p_{21}\,X^2\,Y + p_{12}\,X\,Y^2 + p_{03}\,Y^3 + p_{22}\,X^2\,Y^2 + p_{13}\,X\,Y^3 + p_{04}\,Y^4 \quad \text{(Equation 5c)}$$

poly32
(nine coefficients)

$$sf(x,y) = p_{00} + p_{10}\,X + p_{01}\,Y + p_{20}\,X^2 + p_{11}\,X\,Y + p_{02}\,Y^2 + p_{30}\,X^3 + p_{21}\,X^2\,Y + p_{12}\,X\,Y^2 \quad \text{(Equation 5d)}$$

poly33
(ten coefficients)

$$sf(x,y) = p_{00} + p_{10}\,X + p_{01}\,Y + p_{20}\,X^2 + p_{11}\,X\,Y + p_{02}\,Y^2 + p_{30}\,X^3 + p_{21}\,X^2\,Y + p_{12}\,X\,Y^2 + p_{03}\,Y^3 \quad \text{(Equation 5e)}$$

The relative advantage of poly23 compared to the different alternatives is illustrated in **Figure 3-10**. The figure shows cross-validated RMSE between VIIRS and MODIS FAPAR time series for different models. While the differences between the five polynomials are only relatively small, poly23 performs best and only requires 9 parameters. Indeed, poly24 with twelve parameters is probably overfitting the data, while poly22 is underfitting the data.

Other consolidation stages could not be tested as the necessary cross-validation is extremely time demanding. However, based on the findings shared above, it can be expected that the general trend shown in **Figure 3-10** also applies to other consolidation stages. This makes poly23 the best choice for a 2-dimensional polynomial.

Method	ORIG	DELTA	QMF2	POLY23
RMSEcv	5.30	4.38	4.49	4.20

Polynomial	poly22	poly23	poly24	poly32	poly33
RMSEcv	4.29	4.20	4.23	4.21	4.21

Figure 3-10. *Cross-validated RMSE for different FAPAR intercalibration methods applied to FAPAR consolidation stage C2. Highlighted is the best performing method: poly23. The dataset consists of 324,004 pixels, 36 dekads and 6 years (2018-2023).*

Table 3-5. *Storage requirements (e.g., number of data points to be stored for doing the intercalibration) for a single pixel ("sub-total") as well as the global 500m grid with roughly 570 million pixel ("total") for three competing FAPAR intercalibration methods: offset correction (DELTA), quantile mapping (QMF2) and 2-dimensional polynomial (POLY). The numbers are given for the total of six consolidation stages (C0 to CF) and the 36 dekads per year.*

	DELTA	**QMF2**	**POLY**
parameter	1	101 x 2	9
x dekads	36	36	n.a. (1)
x stages	6	6	6
= sub-total	216	43.632	54
x pixel	570.000.000	570.000.000	570.000.000
= total	$1{,}231 \times 10^{11}$	$2{,}487 \times 10^{13}$	$3{,}078 \times 10^{10}$

As an additional decision support, the storage requirements of various FAPAR intercalibration methods were assessed. As expected – and ignoring ORIG – the storage requirements vary strongly between the three FAPAR intercalibration approaches, with the 2D polynomial (POLY) being the most efficient and the traditional quantile mapping (QMF2) being the worst (**Table 3-5**).

The table shows a significant advantage of DELTA and POLY over the quantile mapping (QMF2) in the order of two to three orders of magnitude. The 2-dimensional polynomial approach (POLY) requires four times less parameter storage than DELTA (e.g., 54 parameters per year instead of 216). This advantage is however associated with a higher computational cost as the offsets in the DELTA approach can be pre-calculated (**Equation 1**), whereas the offsets are varying with the (uncorrected) VIIRS FAPAR value in the case of POLY (**Equation 4**).

On the pure compute side (e.g., in the "calibration phase"), it is obviously computationally cheaper to calculate the offset required for the DELTA approach, compared to fitting for each pixel the coefficients of the 2-dimensional polynomial (POLY):

- For each pixel in the DELTA approach, one needs to load all FAPAR observations of the two sensors for a given dekad, calculate the respective averages and the difference between the averages. This process needs to be repeated 36 times and for each consolidation stage.
- For each pixel in the POLY approach, one has instead to load all FAPAR observations of all dekads of the two sensors, order them separately within each dekad, subtract the ordered FAPAR values and fit the nine coefficients as a function of the respective (uncorrected) VIIRS value and dekad.

4 DISCUSSION OF ALTERNATIVE FAPAR INTERCALIBRATION TECHNIQUES

In the previous section, the alternative FAPAR intercalibration approaches have been compared and quantitatively assessed on a global grid with 324,004 pixels and 36 dekads and 6 validation years (2018 to 2023). In this subsection, we report the results of Task 3: the opportunities and limitation of alternative approaches. For completeness, we involve in the discussion also the baseline of 'no correction' (ORIG). This leads to the conclusions and recommendations in the final subsection.

To allow an easy discussion, the overview in **Table 4-1** has been prepared. The overview specifies different assessment criteria and ranks the alternative intercalibration approaches from 'best' to 'worst'. Also indicated are the supporting figures and tables so that the reader can quickly assess the supporting information.

Table 4-1. *Ranking of three FAPAR intercalibration methods (DELTA, QMF2 and POLY) against the baseline (ORIG) based on cross-validated error assessments involving a global FAPAR dataset of 324,004 pixels, 36 dekads and 6 years (2018-2023). Where available, supporting tables and figures are also indicated*

		Assessment criteria								
		Overall error reduction	Reduction of highest errors	NRT error (C0 and C1) reduction	Error reduction for crop classes	Error reduction for marginal land classes	Seasonal stability	Inter-annual stability	Storage cost	Compute cost
Supporting figures		Figure 3-1, Figure 3-3	Figure 3-5, Figure 3-6	Figure 3-2, Figure 3-3	Figure 3-9	Figure 3-9	Figure 3-5	Figure 3-4	/	/
Supporting tables		Table 3-1	/	Table 3-1	Table 3-4	Table 3-4	Table 3-2, Table 3-3	/	Table 3-5	/
Rank -ing	**Best**	POLY	POLY	POLY	DELTA	POLY	POLY	All equal	ORIG	ORIG
		DELTA	DELTA	DELTA	POLY	DELTA	DELTA		POLY	DELTA
	↓	QMF2	QMF2	QMF2	QMF2	QMF2	QMF2		DELTA	POLY
	Worst	ORIG	ORIG	ORIG	ORIG	ORIG	ORIG		QMF2	QMF2

According to most assessment criteria, the polynomial approach (POLY) is better suited than DELTA as the next most competitive approach (**Table 4-1**). The differences are not enormous (**Table 3-1, Table 3-2**) but significant enough to consider an implementation of this novel and innovative approach.

An implementation of the polynomial approach (POLY) would ensure the best possible global match between FAPAR time series from MODIS and VIIRS and would also ensure that the existing MODIS-derived historical statistics can be used when calculating vegetation anomalies. In this respect, it has also to be highlighted that POLY has the lowest intra-annual seasonality in errors (**Table 3-3**).

In particular for the most relevant NRT analysis involving consolidation stages C0 and C1, POLY provides the lowest (cross-validated) errors amongst all methods and amongst (almost) all land cover classes (**Table 3-4**). Although **Table 4-1** lists DELTA as the best performing approach for the land cover classes GCC (grasses & cereal crops) and BLC (broadleaf crops), a closer inspection of **Table 3-4** reveals that the MAD_{cv} differences between DELTA and POLY for the first two consolidation stages are only minimal (0.02 to 0.08 units in the FAPAR range from 0 to 100). This puts the two approaches almost on par with respect to these two land cover classes.

POLY is also the method which performs the best when the errors between the (uncorrected) VIIRS and MODIS FAPAR time series are the highest (**Figure 3-6**) and generally for the class SAV (savannah) and SHR (shrubs) which are often relevant areas for small-holder farmers, herders and other food-insecure populations (**Table 3-4**).

As vegetation anomalies are calculated based on MODIS-derived historical statistics, POLY's high seasonal consistency is also advantageous (**Figure 3-5** and **Table 3-3**) paired with a significant reduction of the largest errors (**Figure 3-6**) – the latter might potentially lead to significant false positives or false negatives. In this respect, POLY shows again advantages with respect to the other methods.

While the 2-dimensional POLY approach is novel and would require a (partial) reprocessing of the VIIRS data archive, this seems not to be infeasible. Indeed, the calculation of the offset which has to be added to the uncorrected VIIRS FAPAR (**Equation 5a**), is only slightly more demanding compared to DELTA. Storage requirements of POLY are even lower compared to DELTA.

The only more demanding step is the fitting of the polynomials which has to be done separately for each pixel and FAPAR consolidation stage. On a standard PC this "calibration phase" took for the sparse global grid per consolidation stage less than 12 hours, but the process had to be invoked 6 times to abide to the cross-validation requirement.

In a future implementation, however, the whole "calibration" needs only to be run once per pixel and consolidation stage, without any cross-validation. Modern (cloud) computing systems can handle this in a straight forward manner even if the total data amount is (roughly) 2000 x times larger.

5 EXECUTIVE SUMMARY AND RECOMMENDATIONS

The four tasks have been successfully completed. The quantile mapping with a window size of 2 dekads (QMF2) has been implemented (Task 1) and compared with the mean difference correction (DELTA) (Task 2). Similarly, a novel and innovative quantile mapping approach based on 2-dimensional polynomials (POLY) has been implemented and tested (Task 5). The pros and cons of different strategies (Task 3) have been presented and discussed in the previous subsection. Here, we report our final conclusions and recommendations.

First and foremost, an implementation of the QMF2 approach can not be recommended as this approach was in all aspect inferior to both DELTA and POLY (albeit usually better than ORIG). In previous reports good results were reported for QMF2, but overfitting issues were not properly considered. Overfitting issues were avoided in this study thanks to the implemented 6-fold cross-validation.

The DELTA approach has the advantage of being already implemented and used. Also, compared to POLY, results are not dramatically worse. However, compared to POLY several disadvantages need to be highlighted and considered (**Table 4-1**), and here most importantly the lower performance for the relevant early consolidation stages (C0 and C1) and the inferior performance under situations were VIIRS and MODIS are the least well aligned. POLY clearly shows the most consistent spatial and temporal performance, unmatched by the other approaches. This ensures a globally homogeneous dataset with minimal seasonal and land cover related effects.

Based on the above presented findings, and despite the somehow higher computational costs of POLY, the implementation of the 2-dimensional polynomial approach (POLY) is strongly recommended.

6 ANNEX: SUPPLEMENTARY RESULTS

For the 70 locations indicated in **Figure 2-1**, the analysis had been done separately. The results are reported here for completeness, although they do not not differ materially from the global analysis presented in the main body of the report (e.g., identical methods were used).

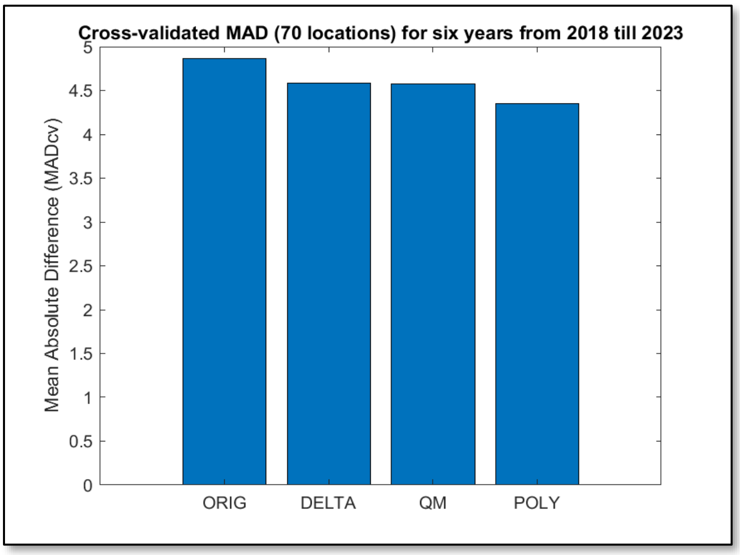


Figure 6-1. *Cross-validated mean absolute differences (MADcv) for the six validation years 2018-2023 and the three analyzed intercalibration methods (DELTA to POLY). For comparison, the baseline method (ORIG) is also added. The analysis is based on FAPAR data from the C0 stage and involves FAPAR data from 70 pixels, 36 dekads and 6 years. For the 70 locations, please refer to Figure 2-1.*

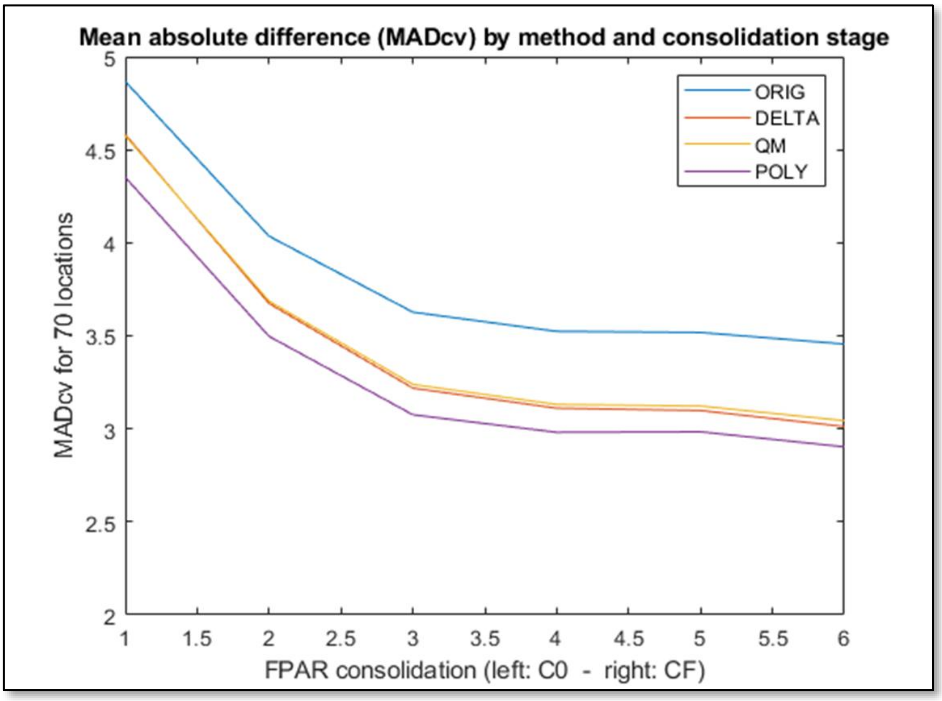
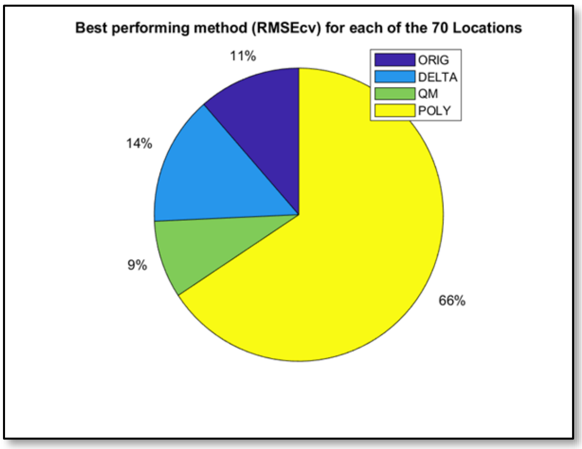


Figure 6-2. *Impact of the consolidation stage (from C0 left to CF on the right) on the cross-validated mean absolute differences (MADcv) for the six validation years 2018-2023. The analysis is based on FAPAR data from 70 pixels, 36 dekads and 6 years. The three intercalibration methods are DELTA, QMF2 and POLY. The uncorrected baseline is ORIG. The reader is referred to Figure 2-1 for the location of the 70 sample points.*



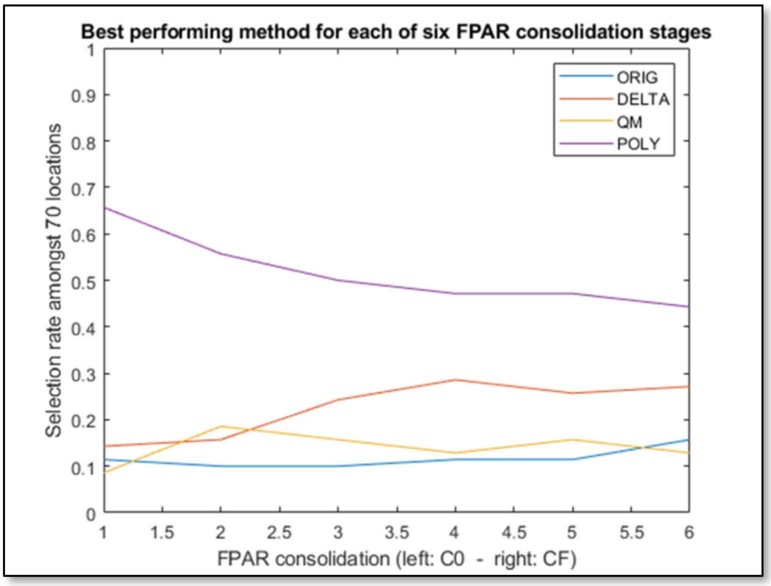


Figure 6-3. *Relative frequency with which each method leads to the lowest mismatch between VIIRS and MODIS FAPAR based on the cross-validated RMSE. (top) pie-chart for FAPAR consolidation stage C0. (bottom) Selection rates for all 6 FAPAR consolidation stages and methods (from C0 left to CF right). The results are based on the analysis of 70 pixels, 36 dekads and 6 validation years.*

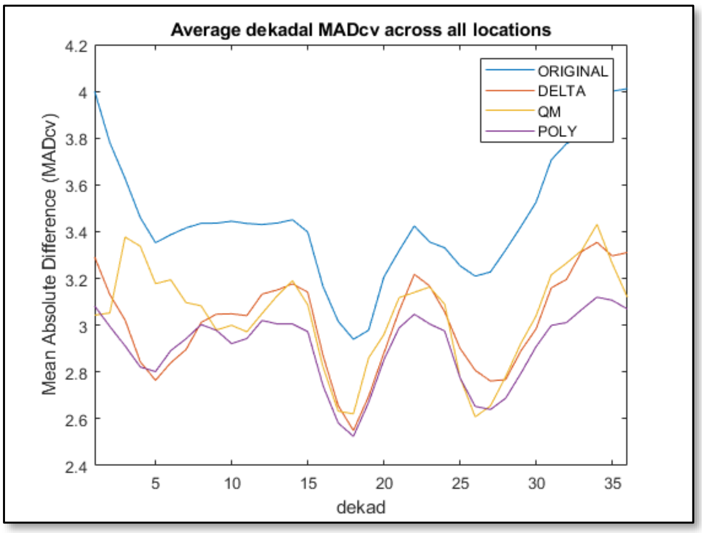


Figure 6-4. *Cross-validated RMSE between VIIRS and MODIS FAPAR by dekad for the CF consolidation stage and the six validation years 2018-2023. Per dekad, the analysis is based on FAPAR data from 70 pixels and 6 years. The three FAPAR intercalibration methods are DELTA, QMF2 and POLY. The uncorrected baseline is ORIG.*

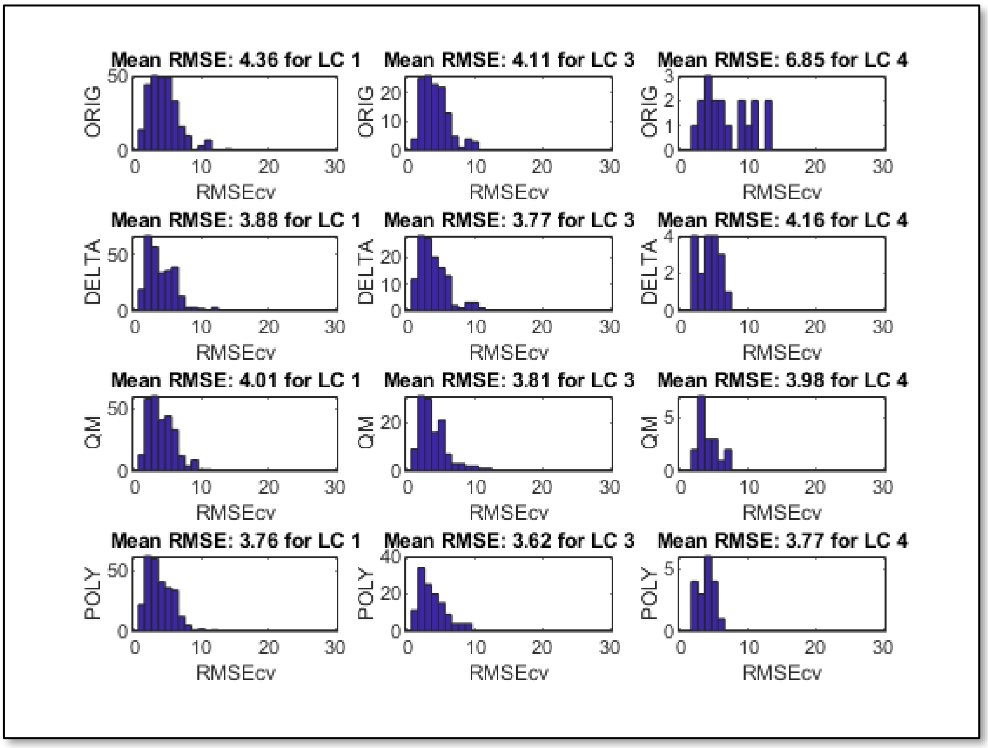


Figure 6-5*. Cross-validated RMSE between VIIRS and MODIS FAPAR for three land cover classes and the six validation years 2018-2023 (for C0). The three FAPAR intercalibration methods are DELTA, QMF2 and POLY. The uncorrected baseline is ORIG. The three land cover classes are: LC1 – Grasses/Cereal Crops; LC3 – Broadleaf Crops; LC4 – Savannah. Results are for the 70 locations shown in* ***Figure 2-1***

The **Figure 6-1 to Figure 6-5** re-confirm the findings presented in the main body of the report. The following statements can be reiterated based on this small subsample of 70 European and African locations:

- There is no real justification for using the computationally much more demanding QMF2 approach, compared to the much simpler offset correction (DELTA)
- While QMF2 and DELTA achieve similar error reductions, and this across the MODIS consolidation stages from C0 to CF, the best results are obtained using the 2-dimensional polynomial approach (POLY)
- POLY performs best across dekads, land cover types and consolidation stages

Together, these findings reaffirm the recommendation for implementing the POLY approach to best intercalibrate FAPAR time series from MODIS and VIIRS.