Copernicus FICE 2025

Training on
In situ Ocean Colour Above-Water Radiometry towards Satellite Validation

Application of field data to assess satellite data products. Comparing performance metrics with different matchup protocols

Vittorio Brando
CNR-ISMAR
Vittorio.brando@cnr.it

PROGRAMME OF THE EUROPEAN UNION
Copernicus Europe's eyes on Earth
IMPLEMENTED BY EUMETSAT
FRM4SOC Phase-2
fiducial reference measurements for satellite ocean colour
CNR ISMAR ISTITUTO DI SCIENZE MARINE
Venice International University

6-20 July 2025
Venice, Italy

CNR ISMAR
ISTITUTO DI SCIENZE MARINE

Remote Sensing of Environment 259 (2021) 112415

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

ELSEVIER

Check for updates

Assessing the influence of different validation protocols on Ocean Colour match-up analyses

Javier A. Concha*, Marco Bracaglia, Vittorio E. Brando

Istituto di Scienze Marine, Consiglio Nazionale delle Ricerche, Via Fosso del Cavaliere 100, 00133 Rome, Italy

frontiers | Frontiers in Remote Sensing

TYPE Original Research
PUBLISHED 18 March 2024
DOI 10.3389/frsen.2024.1330317

Check for updates

OPEN ACCESS

EDITED BY
Enner Alcântara,
São Paulo State University, Brazil

REVIEWED BY
Yulong Guo,
Henan Agricultural University, China
Shun Bi,
Helmholtz Association of German Research Centres (HZ), Germany

*CORRESPONDENCE

Validation of satellite water products based on HYPERNETS *in situ* data using a Match-up Database (MDB) file structure

Luis González Vilas[1]*, Vittorio E. Brando[1], Javier A. Concha[1,2], Clèmence Goyens[3], Ana I. Dogliotti[4,5], David Doxaran[6], Antoine Dille[3] and Dimitry Van der Zande[3]

COPERNICUS MARINE ENVIRONMENT MONITORING SERVICE
CMEMS OCEAN COLOR TAC

OCTAC CMEMS

HYPERNETS

CONSIGLIO NAZIONALE DELLE RICERCHE
ISTITUTO DI SCIENZE MARINE

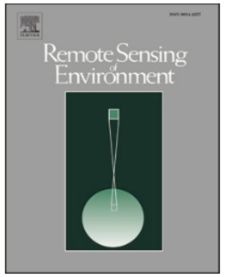
CNR ISMAR
ISTITUTO DI SCIENZE MARINE

Remote Sensing of Environment 259 (2021) 112415

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse



Assessing the influence of different validation protocols on Ocean Colour match-up analyses

Javier A. Concha *, Marco Bracaglia, Vittorio E. Brando

Istituto di Scienze Marine, Consiglio Nazionale delle Ricerche, Via Fosso del Cavaliere 100, 00133 Rome, Italy

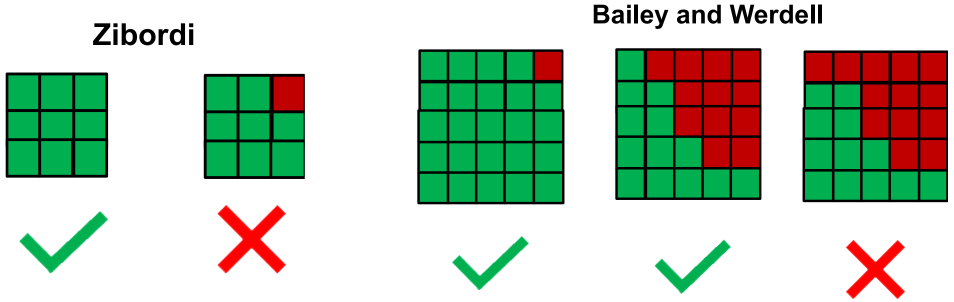
OCTAC CMEMS

HYPERNETS

This community uses validation protocols based mainly in two approaches:

Authors	Minimum number of valid pixels within the extract to be considered	CV criteria, bands used for CV, CV threshold (spatial homogeneity test)	SZA, VZA thresholds	Number of elements in satellite extract, mean used	Temporal window
Z09 **(Zibordi et al 2009)**	**100% (9 pixels)**	Both Lwn(555) <AND> AOT(865), 0.2 (20%)	70°, 56°	**3x3, average** (Zibordi et al. 2009), statistic used	±2 hr
BW06 **(Bailey and Werdell, 2006)**	**50%+1 (13 pixels)**	Median of CV of 412-555 nm and AOT(865), 0.15 (15%)	75°, 60°	**5x5, filtered mean** Filtering: Value is within +/- 1.5 *sd plus mean	±3 hr

CV: Coeff. Of Variation CV = SD/mean
AOT: Aerosol Optical Thickness
SZA: solar zenith angle
VZA: viewing zenith angle



Remote Sensing of Environment 259 (2021) 112415

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Assessing the influence of different validation protocols on Ocean Colour match-up analyses

Javier A. Concha*, Marco Bracaglia, Vittorio E. Brando

Istituto di Scienze Marine, Consiglio Nazionale delle Ricerche, Via Fosso del Cavaliere 100, 00133 Rome, Italy

Table 2
Quality check and homogenization steps implemented in different validation studies. ND: not determined. N/A: not applicable.

Authors	Minimum number of valid pixels within the extract to be considered	Flagging	CV criteria, bands used for CV, CV threshold (spatial homogeneity test)	SZA, VZA thresholds
Zibordi et al. (2009a)	100% (9 pixels)	Standard flags of the SeaDAS processing code (cloud, sunglint, high solar and view zenith angles)	Both $L_{WN}(555)$ <AND > AOT(865), 0.2 (20%)	70°, 60°
Zibordi et al. (2018)	100% (9 pixels)	Not affected by cloud contamination, and in general, by any of the main exclusion flags, EXCEPT ANNOT flags	Both $L_{WN}(555)$ <AND > AOT(865), 0.2 (20%)	70°, 56°
Bailey and Werdell (2006)	50%. For coastal, a minimum of 50% of the non-land pixels, with an absolute minimum of 5 pixels	Land, cloud or ice, stray light, sun glint, high TOA radiance, low L_{WN} (555), or atmospheric correction failure	Median of CV of 412–555 nm and AOT(865), 0.15 (15%)	75°, 60°
Volpe et al. (2019)	≥ 5 pixels,	L2 flags provided by space agencies: OBPG except ATMFAIL for VIIRS	0.2 (20%); To which product or products this is applied is not mentioned	75°, 60
Qin et al. (2017)	≥ 5 pixels,	Land, haze, whitecaps, cloud or sun glint contamination flags based on processing the L1B data with Idepix v2.2.10. Processor specific flags included: poor fits to aerosol models; TOA radiances outside of the training or application range; and results surpassing the minimum or maximum concentration bounds.	Same as Bailey and Werdell (2006) = 0.15 (15%)	ND
Cui et al. (2010)	Same as Bailey and Werdell (2006)	Flags MERIS L2, such as sunglint	Similar as Bailey and Werdell (2006), **but with CV calculated with the median instead of the mean.**	ND
Müller et al. (2015)	≥ 5 pixels,	Flags from specific AC alg.	3-sigma filter for outliers, CV but with the median instead of the mean < 0.15	N/A (depends AC processor)
Barnes et al. (2019)	extract discarded if ≥4 of the 9 pixels in the 3 × 3 pixels box are flagged	AC L2 flags	various and 0.2 (20%)	ND
Vanhellemont (2019)	100%	Not used (Thresholds to specific band used instead)	Not used (Thresholds to specific band used instead)	ND
Ilori et al. (2019)	Same as Bailey and Werdell (2006)	land, clouds, cloud-shadow, ice, stray light, low $L_{WN}(555)$, high VZA (> 60), high sunglint, and high TOA radiance	Same as Bailey and Werdell (2006)	Same as Bailey and Werdell (2006)
Van der Zande et al. (2016)	1 pixel	Not used (Thresholds to specific band used instead)	N/A (central pixel used)	ND
Warren et al. (2019)	N/A	invalid, cloud, cloud_ambiguous, cloud_sure, cloud_buffer, cloud_shadow, cirrus_sure, cirrus_ambiguous, land and vegrisk	N/A (central pixel used)	Same as Bailey and Werdell (2006)
Caballero et al. (2018)	100% (9 pixels) (not mentioned explicitly)	POLYMER flags for masking pixels with sunglint	Not mentioned	ND
Pahlevan et al. (2018)	Not mentioned	hazy images and partial sunglint	Not mentioned	ND

Remote Sensing of Environment 259 (2021) 112415

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

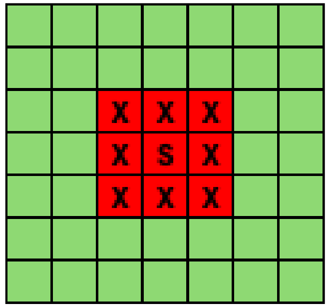
Assessing the influence of different validation protocols on Ocean Colour match-up analyses

Javier A. Concha *, Marco Bracaglia, Vittorio E. Brando

Istituto di Scienze Marine, Consiglio Nazionale delle Ricerche, Via Fosso del Cavaliere 100, 00133 Rome, Italy

Table 3
Spatiotemporal collocation choices implemented in different validation studies.

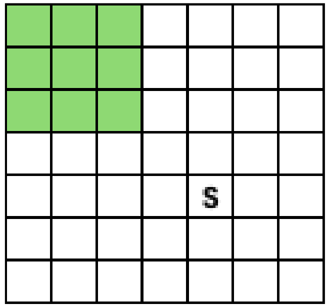
Authors	Number of elements in satellite extract, quantity reported	Temporal window
Zibordi et al. (2009a)	3 × 3, **mean**	±2 h
Zibordi et al. (2018)	3 × 3, **mean**	±2 h (70% ±1 h) for AERONET-OC, ±4 h for BioMaP (50% ±2 h)
Bailey and Werdell (2006)	5 × 5, **filtered mean** Bailey and Werdell (2006), Filtering: Value is within +/− 1.5 *standard deviation plus mean, with standard deviation evaluated on the x*x pixels window.	±3 h for homogenous water masses
Volpe et al. (2019)	3 × 3; **median**; 1 km resolution on the equirectangular grid covering	Same day for L3
Qin et al. (2017)	3 × 3; **mean** of the pixel that pass: differences between the value of each valid pixel and their mean in the box were limited to twice the standard deviation to eliminate outliers	±12 h or ± 3 h (ship) and ± 2 h (AERONET-OC); sensitivity study for ±0.5, ±2, ±3, ± 4, ±6 and ± 12 h
Cui et al. (2010)	3 × 3; Similar as Bailey and Werdell (2006) but with **median** of filtered and valid values, instead of mean. Filtering: Value is within +/− 1.5 *standard deviation plus **median**, with standard deviation evaluated on the x*x pixels window.	Same as Bailey and Werdell (2006)
Müller et al. (2015)	3 × 3, median and standard deviation	±3 h
Barnes et al. (2019)	Mentions: 3 × 3, 5 × 5, 11 × 11; Used: 3 × 3, mean	Mentions: ±2, ±3, ±3.5, ±8 h; Used: ±2 h
Vanhellemont (2019)	3 × 3 (manually shifted), **mean**	±2 h
Ilori et al. (2019)	7 × 7, removing 3 × 3 to avoid the structure and shadow	±0.5 h
Van der Zande et al. (2016)	1 × 1	±1 h
Warren et al. (2019)	ONLY central pixel was used for the regression; 3 × 3 at 60 m (180 × 180 m) was used for Deming regression	±3 h for Coastal, ±24 h for inland
Caballero et al. (2018)	5 × 5, **mean**	<±0.5 h; <±1 h
Pahlevan et al. (2018)	Not mentioned for AERONET-OC	Not mentioned



For Landsat & Sentinel 2 imagery 2 at 10/30 m resolution, the structure (ship or tower) may be large enough to perturb the signal in the satellite imagery

"the doughnut" (Ilori et al, 2019)

"manual shift" (Vanhellemont, 2019)



Which one to choose? Why?

Authors	Minimum number of valid pixels within the extract to be considered	CV criteria, bands used for CV, CV threshold (spatial homogeneity test)	SZA, VZA thresholds	Number of elements in satellite extract, mean used	Temporal window
Z09 **(Zibordi et al 2009)**	**100% (9 pixels)**	Both Lwn(555) <AND> AOT(865), 0.2 (20%)	70°, 56°	**3x3, average** (Zibordi et al. 2009), statistic used	±2 hours
BW06 **(Bailey and Werdell, 2006)**	**50%+1 (13 pixels)**	Median of CV of 412-555 nm and AOT(865), 0.15 (15%)	75°, 60°	**5x5, filtered mean** Filtering: Value is within +/- 1.5 *sd plus mean	±3 hours

Goal

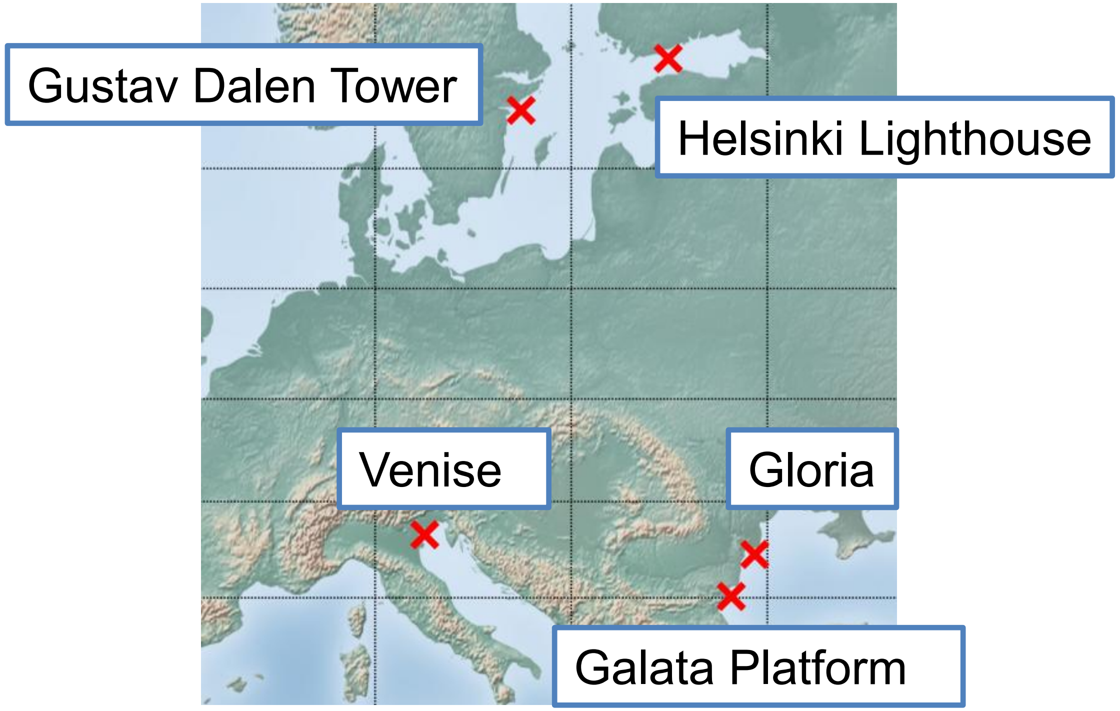
- The aim of this work will be the comparison of the effects of the differences between the methods to better inform the selection of validation variants:
  - **Z09:** Zibordi et al., RSEnv (2009).
  - **BW06:** Bailey and Werdell, RSEnv (2006).
- For medium spatial resolution (S3A/OLCI)
- For the same in situ dataset

Note: This is not an accuracy assessment of S3A with specific interpretation of the match-up analyses for sites, water types or processing algorithms.

In situ data:

AERONET-OC

- Level 1.5 and 2.0
- 5 stations in Mediterranean, Baltic and Black Sea



OLCI dataset:

- Normalized Water-leaving radiance from Sentinel-3A/OLCI Full Resolution (WFR, 300 m) Level 2 data for IPF-OL-2 version 6.13:

$$L_{\mathrm{WN}}^{\mathrm{OLCI}}(\lambda) = \rho^{OLCI}(\lambda)\frac{F0(\lambda)}{\pi}C_{f/Q}(\lambda)$$

where:

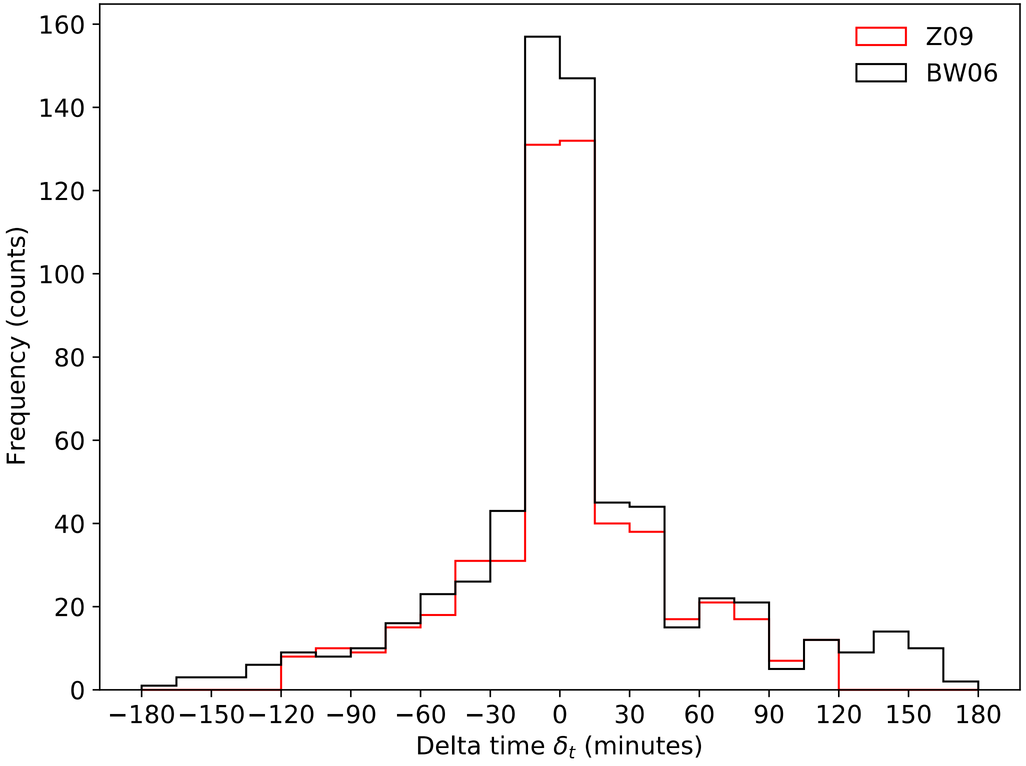
- ρ^{OLCI}: spectral reflectance
- $F0$: the mean extraterrestrial solar irradiance
- $C_{f/Q}$: is the BRDF correction factor (Morel, Antoine, and Gentili 2002).

Number of Matchups

- Different total number of match-ups for the validation protocols:
  BW06 produces 20% more match-ups.
- Increasing or decreasing depending station.

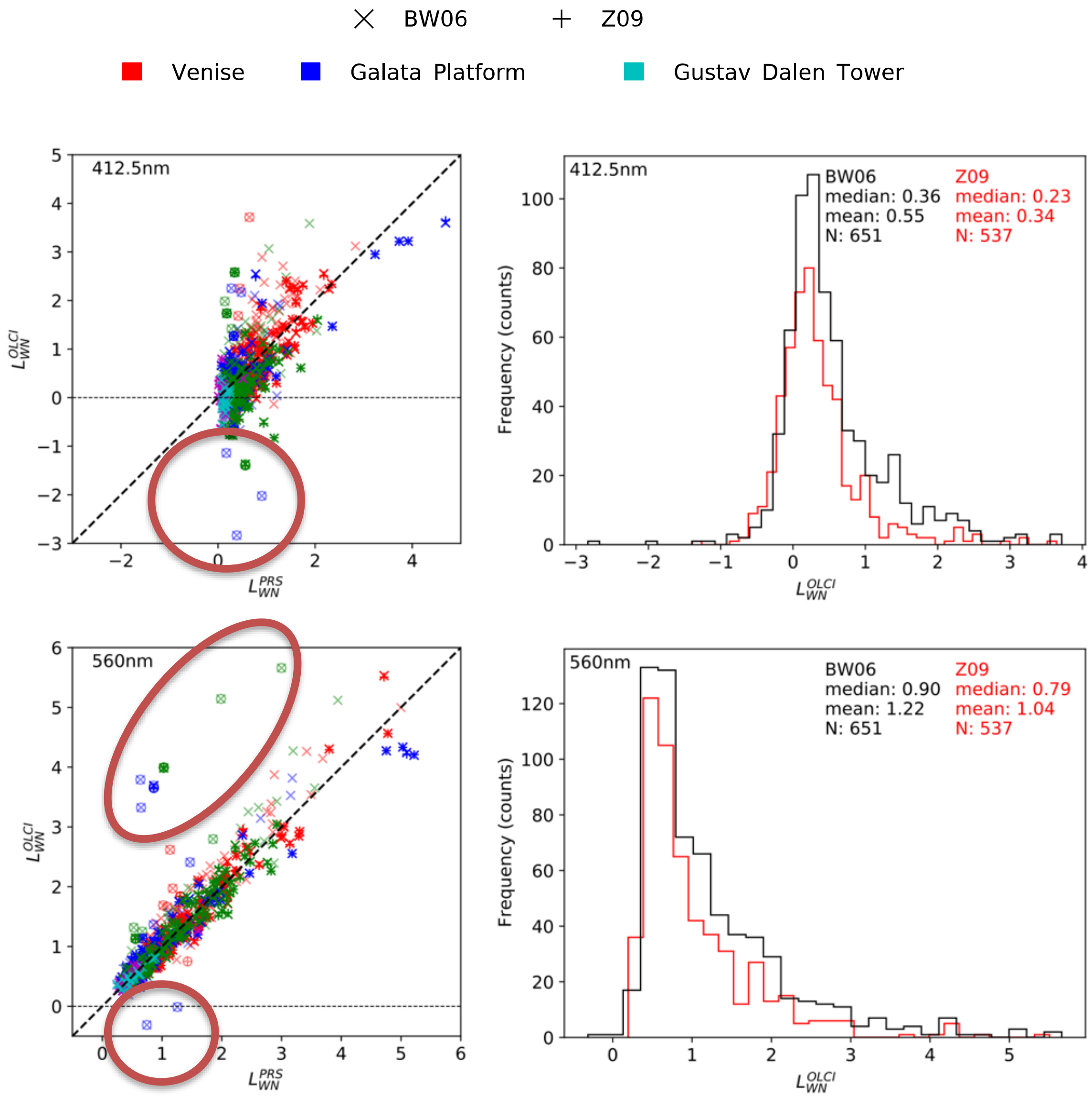
	Potential MUs		MUs			
			BW06		Z09	
Station	$\Delta_t = 2h$	$\Delta_t = 3h$	$\Delta_t = 2h$	$\Delta_t = 3h$	$\Delta_t = 2h$	$\Delta_t = 3h$
Venise	313	378	188	**211**	**137**	154
Gloria	340	374	199	**211**	**184**	194
Galata Platform	384	403	172	**179**	**150**	158
Helsinki Lighthouse	76	93	22	**23**	**36**	40
Gustav Dalen Tower	47	66	22	**27**	**30**	41
N_{total}	1160	1314	603	**651**	**537**	587

δ_t: Sat. time – In situ time



- 80% of match-ups occur within 1 hour for both BW06 and Z09.

Let’s look at the matchups



Distribution of satellite data

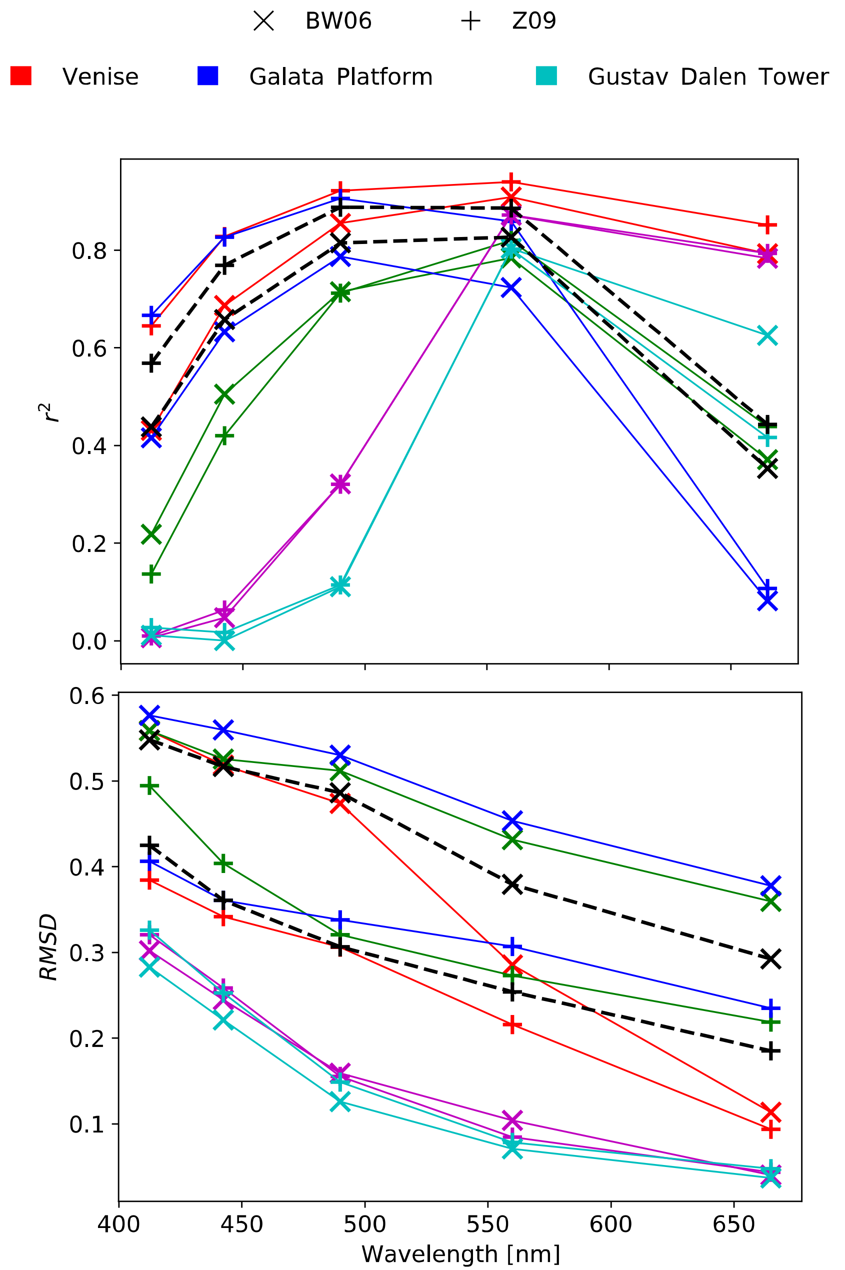
Protocol	λ[nm]	min	max	std	median	mean	N
BW06	412.5	-2.84	3.71	0.73	0.36	0.55	651
Z09	412.5	-1.39	3.64	0.58	0.23	0.34	537
BW06	442.5	-2.52	5.36	0.88	0.56	0.84	651
Z09	442.5	-0.51	5.39	0.73	0.43	0.61	537
BW06	490.0	-1.57	7.03	1.07	0.92	1.28	651
Z09	490.0	-0.11	7.03	0.91	0.76	1.02	537
BW06	560.0	-0.31	5.66	0.90	0.90	1.22	651
Z09	560.0	0.20	5.51	0.74	0.79	1.04	537
BW06	665.0	-0.46	3.27	0.36	0.13	0.22	651
Z09	665.0	-0.11	3.15	0.25	0.11	0.17	537

- Median and mean values different for the different protocols.
- More outliers pass BW06.
- More negatives values pass the BW06.

Validation Metrics

Protocol	λ[nm]	N	RMSD	MAPD[%]	MPD[%]	MB	MAD	r^2
BW06	412.5	651	0.55	105.9	-1.7	-0.05	0.39	0.44
Z09	412.5	537	0.42	105.7	-45.5	-0.18	0.33	0.57
BW06	442.5	651	0.52	61.2	6.1	0.04	0.34	0.66
Z09	442.5	537	0.36	59.4	-18.6	-0.09	0.26	0.77
BW06	490.0	651	0.49	30.0	12.2	0.13	0.28	0.81
Z09	490.0	537	0.31	25.2	1.0	0.01	0.19	0.89
BW06	560.0	651	0.38	18.9	5.4	0.05	0.20	0.83
Z09	560.0	537	0.25	14.6	-1.0	-0.02	0.14	0.89
BW06	665.0	651	0.29	69.2	16.5	0.02	0.09	0.35
Z09	665.0	537	0.19	46.2	-14.0	-0.02	0.06	0.44

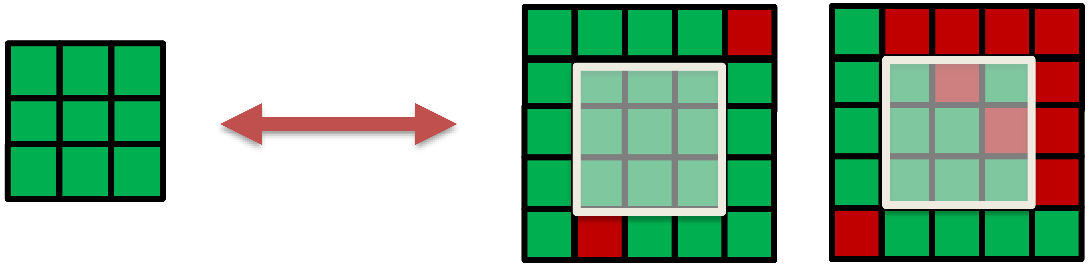
- **BW06 and Z09 metrics are different.**
- Metrics differ by stations
- BW06 produces more match-ups.
- For the total of match-ups, Z09 has better metrics.



Common match-ups

Protocol	λ[nm]	N	RMSD	MAPD[%]	MPD[%]	MB	MAD	r^2
BW06	412.5	451	0.43	95.9	-30.9	-0.1748	0.336	0.56
Z09	412.5	451	0.42	93.9	-27.8	-0.1633	0.325	0.56
BW06	442.5	451	0.37	53.3	-10.1	-0.0731	0.264	0.76
Z09	442.5	451	0.36	52.8	-8.3	-0.0642	0.257	0.76
BW06	490.0	451	0.32	23.4	3.9	0.0285	0.192	0.88
Z09	490.0	451	0.31	23.4	4.7	0.0344	0.190	0.88
BW06	560.0	451	0.27	14.4	-0.1	-0.0158	0.146	0.88
Z09	560.0	451	0.26	14.3	0.1	-0.0143	0.144	0.88
BW06	665.0	451	0.20	45.9	-10.2	-0.0132	0.062	0.42
Z09	665.0	451	0.19	45.0	-9.2	-0.0125	0.061	0.42

- Similar metrics values.
- Main differences in MPD and MB.
- 7 to 9 Z09 pixels are included in 85-90% of the BW06 matchups.



NON Common match-ups

BW06 and not Z09:

N match-ups: 200

Rejection causes:

- 2-h time window: 48
- CV>20%: 57
- 100% valid pixels: 107

Z09 and not BW06:

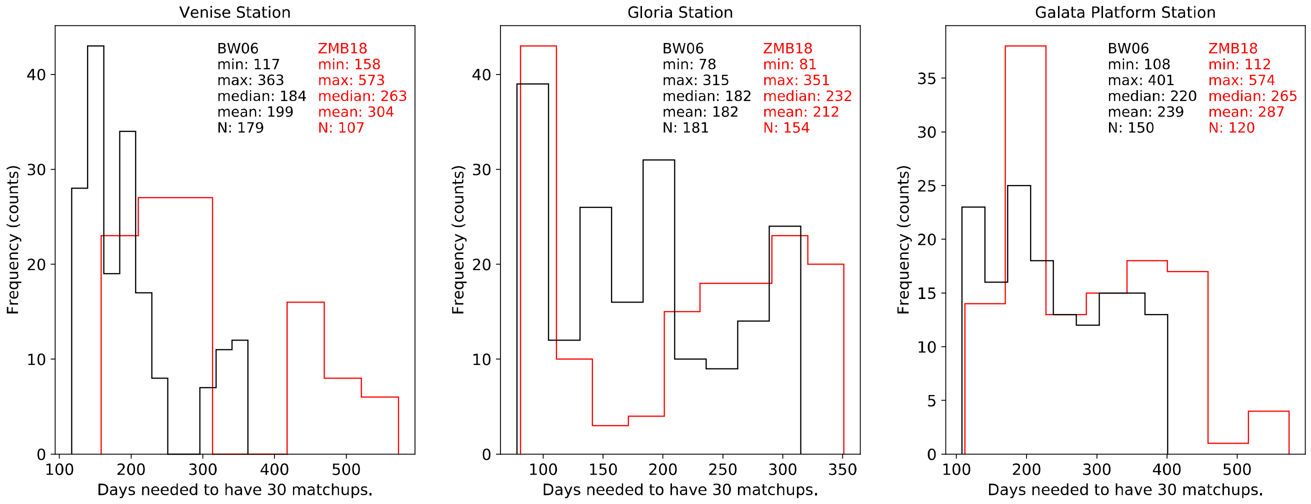
N match-ups: 86

Rejection cause:

- CV>15%: 86

Effects in satellite calibration

Station	Protocol (days)	mean (days)	median (days)	Diff. median (days)
Venise	Z09	303	263	–
Venise	BW06	199	184	79
Galata Platform	Z09	286	265	–
Galata Platform	BW06	238	220	45
Gloria	Z09	212	231	–
Gloria	BW06	181	182	49



- Minimal match-ups for calculation vicarious calibration gains.
- How many days needed to reach 30 match-ups?
- A moving 30-match-up window was applied and the i^{th} and i^{th}+30 days was recorded.
- The number of match-ups needed are reached in a shorter time with BW06 (2.6 months for Venise, 1.5 month for Galata Platform and Gloria).

Environmental Perturbations (u_{en}^{Z09} and u_{en}^{BW06})

Station	MU Dataset	BW06 $Median[CV]$	BW06 $L_{WN}(560)$	Z09 $L_{WN}(560)$	Z09 $AOT(865.5)$
Venise	All	5.8	4.4	4.6	5.3
Galata Platform	All	7.2	4.9	5.5	5.9
Gloria	All	8.4	3.5	5.1	8.0
Helsinki Lighthouse	All	10.9	6.6	8.0	8.2
Gustav Dalen Tower	All	11.2	4.6	5.8	9.2
Venise	Common	5.9	4.3	4.4	5.2
Galata Platform	Common	7.3	4.9	5.2	5.8
Gloria	Common	8.6	3.7	4.2	7.9
Helsinki Lighthouse	Common	10.7	6.9	7.8	8.0
Gustav Dalen Tower	Common	11.0	4.3	5.3	8.7

- Median of the CV were considered as a proxy of the uncertainties due to environmental perturbations in the satellite imagery u_{en}^{BW06} and u_{en}^{Z09}.
- For $L_{WN}(560)$ => u_{en}^{Z09}: 4.2-7.8% and u_{en}^{BW06}: 3.7-6.9%
- For the Baltic Sea: Median[CV] ~ 10%

Conclusions

BW06 produces **more** total of matchups, spanning a wider dynamic range while **Z09** provides **lower** uncertainties figures in most of the validation metrics.

- The number of matchups and metrics depend on the quality checking and spatiotemporal criteria of the protocols.
- Because the high AERONET-OC sampling frequency, most of the match-ups occurs within 1 hour (~80%) and ~60% within 30 min.
- Larger difference between BW06 and Z09 is brought by different quality checking criteria and not the different time windows.

Conclusions

BW06 produces **more** total of matchups, spanning a wider dynamic range while **Z09** provides **lower** uncertainties figures in most of the validation metrics.

- Although the same reference dataset was used, the differences between methods provide a different ”impression of accuracy”.
- For the common match-ups, validation metrics are similar because most of the Z09 pixels are included in the calculation of the BW06 filtered mean for 85% to 90% of the cases.

<u>The accuracy reported in different studies may not always be directly comparable.</u>

CNR ISMAR
ISTITUTO DI SCIENZE MARINE

frontiers | Frontiers in Remote Sensing

TYPE Original Research
PUBLISHED 18 March 2024
DOI 10.3389/frsen.2024.1330317

Check for updates

OPEN ACCESS

EDITED BY
Enner Alcântara,
São Paulo State University, Brazil

REVIEWED BY
Yulong Guo,
Henan Agricultural University, China
Shun Bi,
Helmholtz Association of German Research Centres (HZ), Germany

*CORRESPONDENCE

Validation of satellite water products based on HYPERNETS *in situ* data using a Match-up Database (MDB) file structure

Luis González Vilas[1*], Vittorio E. Brando[1], Javier A. Concha[1,2], Clèmence Goyens[3], Ana I. Dogliotti[4,5], David Doxaran[6], Antoine Dille[3] and Dimitry Van der Zande[3]

CMEMS OCEAN COLOR TAC

OCTAC CMEMS

HYPERNETS

TABLE 2 WATERHYPERNET sites included in the Sentinel-3 validation analysis.

Site	Site code	Country	Water type	Location	Installation date
Berre	BEFR	France	Inland—productive and turbid	43°28′09″ N 5°05′03″ E	2021-02-24
AAOT	VEIT	Italy	Moderately to turbid coastal waters	45°18′51.29″ N 12°30′29.70″ E	2021-04-16
MAGIR	MAFR	France	Estuarine turbid to highly turbid	45°32′43.69″ N 1°02′24.62″ W	2021-11-08
RdP-EsNM	LPAR	Argentina	Estuarine highly turbid	34°49′4.76″ S 57°53′45.28″ W	2021-12-14
Lake Garda	GAIT	Italy	Inland—clear waters (macrophytes)	45°34′35.93″ N 10°34′47.80″ E	2022-06-08
Zeebrugge	M1BE	Belgium	Marine—very turbid	51°21′43.2″ N 3°07′12″ E	2022-11-22

TABLE 3 Default and site-specific protocols implemented in the OLCI WFR validation analysis. Superscripts in the flag list indicate the flags used in the flag groups shown in Figure 3 (1: S3_CLOUD; 2: S3_RWNEG; 3: S3_INVALID; 4: HIGHGLINT; 5: HISOLZEN).

Default
Wavelength (in nm): 400, 412.5, 442.5, 490, 510, 560, 620, 665, 673.75, 681.25, 708.75, 753.75, 778.75,865, 885
Measurement window size: 3 × 3 pixels
Minimum number of valid pixels: 9
Flag List (WQSF): LAND, COASTLINE, CLOUD[1], CLOUD_AMBIGUOUS[1], CLOUD_MARGIN[1], RWNEG_O2[2], RWNEG_O3[2], RWNEG_O4[2], RWNEG_O5[2], RWNEG_O6[2], RWNEG_O7[2], RWNEG_O8[2], INVALID[3], AC_FAIL[3], SUSPECT[3], HIGHGLINT[4], HISOLZEN[5], COSMETIC, SATURATED, SNOW_ICE, WHITECAPS
Reported quantity: Average after excluding outliers
Geometry thresholds
Solar Zenith Angle (SZA) > 70°
Viewing Zenith Angle (OZA) > 70°
Spatial Homogeneity Test: CV > 20% at 560 nm
Time window: 2 h

Site-specific protocols	
BEFR	Default
	Masked pixels with negative Rrs at 400 nm, 412.5 nm or 442.5 nm
VEIT	Default
MAFR	Default
	Minimum number of valid pixels: 1
	Masked pixels with negative Rrs at 442.5 nm
	NIR similarity spectrum correction is not applied for in situ data
LPAR	Default
	Masked pixels with negative Rrs at 442.5 nm
	NIR similarity spectrum correction is not applied for in situ data
GAIT	Default
	Minimum number of valid pixels: 1
M1BE	Default
	NIR similarity spectrum correction is not applied for in situ data

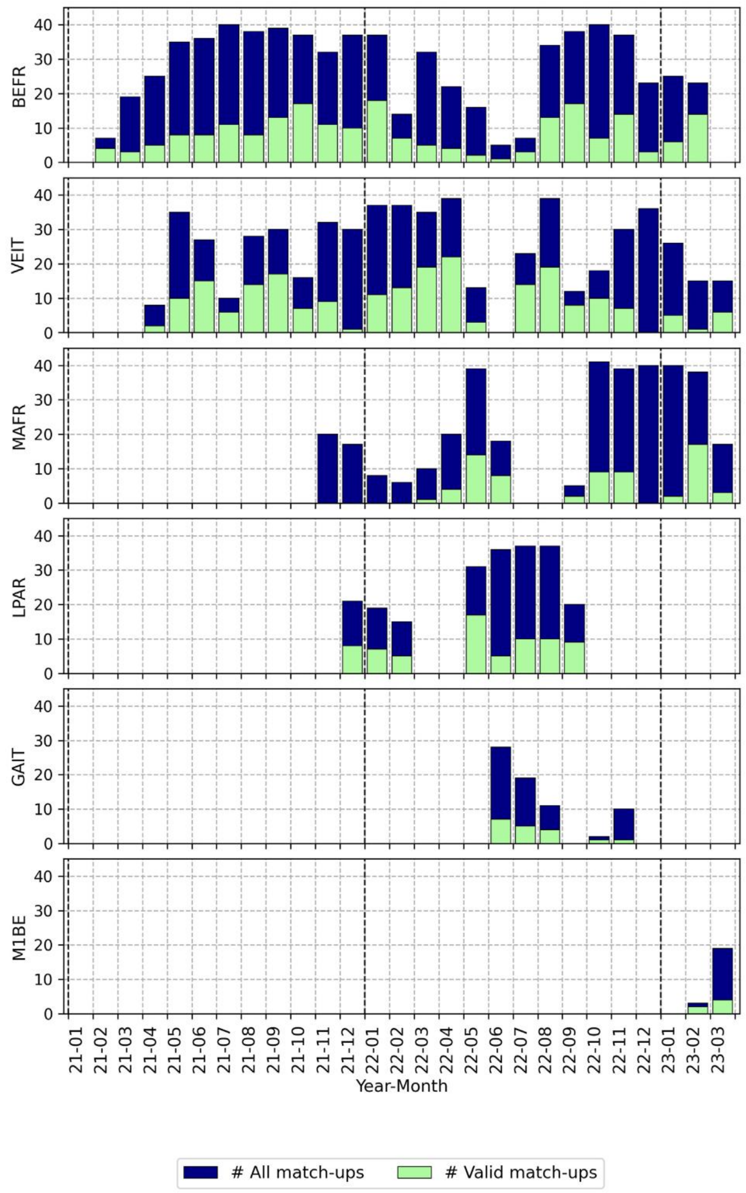


FIGURE 2
Number of total and valid Sentinel-3 WFR match-ups per month and site.

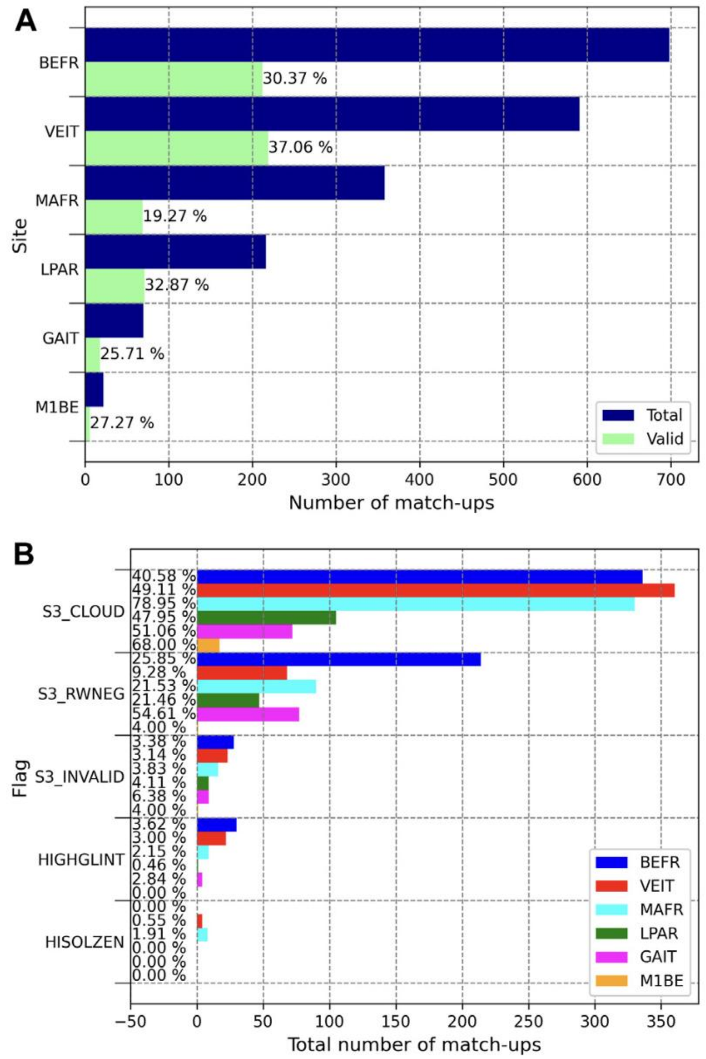


FIGURE 3
(A) Number of total and valid match-ups for each site. Validity rate is also shown. **(B)** Number of match-ups with at least one pixel affected by a specific flag within the measurement window, and percentage with respect to the total number of potential match-ups. Flags included in each flag group are shown in Table 1.

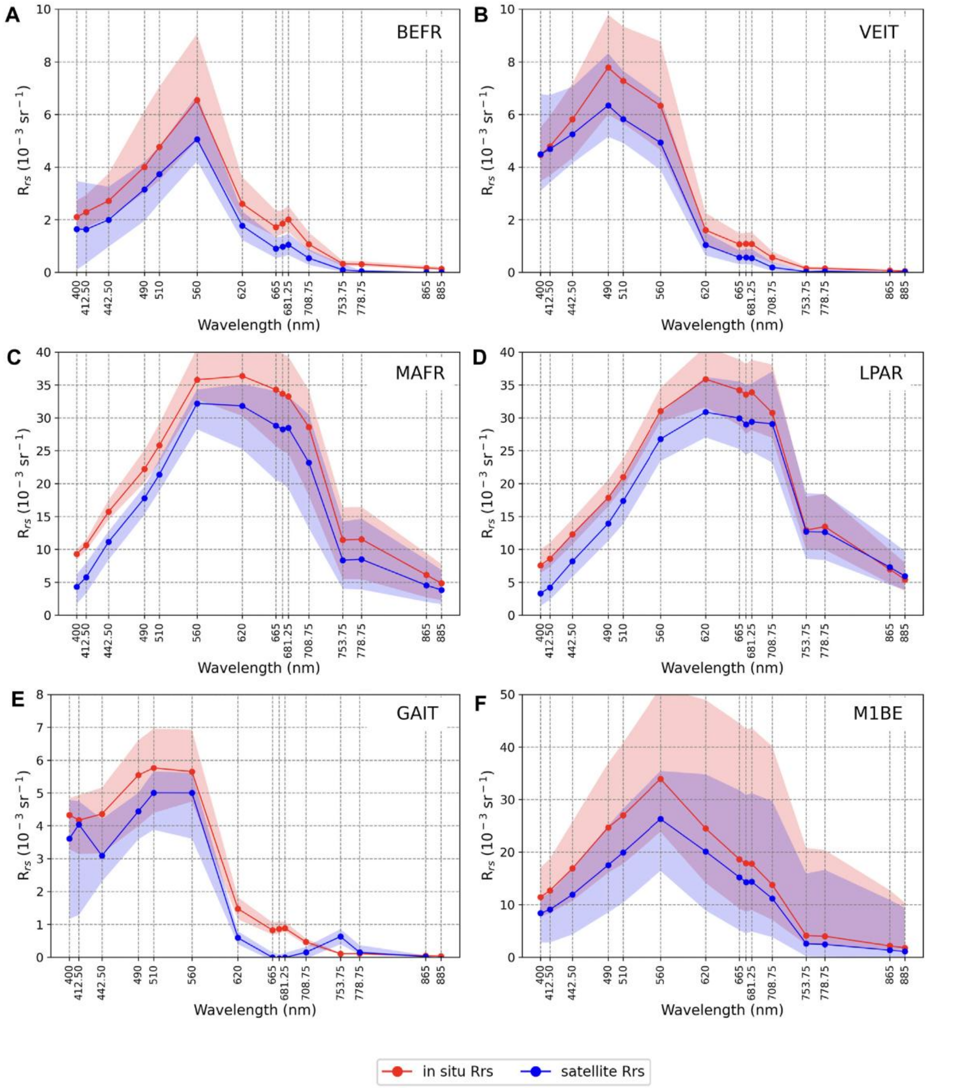


FIGURE 4
Comparison between satellite (Sentinel-3 WFR) and *in situ* (HYPSTAR® L2) spectra for each site: **(A)** BEFR; **(B)** VEIT; **(C)** MAFR; **(D)** LPAR; **(E)** GAIT; **(F)** M1BE. Lines show the average spectra and shadow areas indicate the interquartile range. Note that spectral shapes and Rrs ranges reveal, to some extent, the specific optical characteristics of each site.

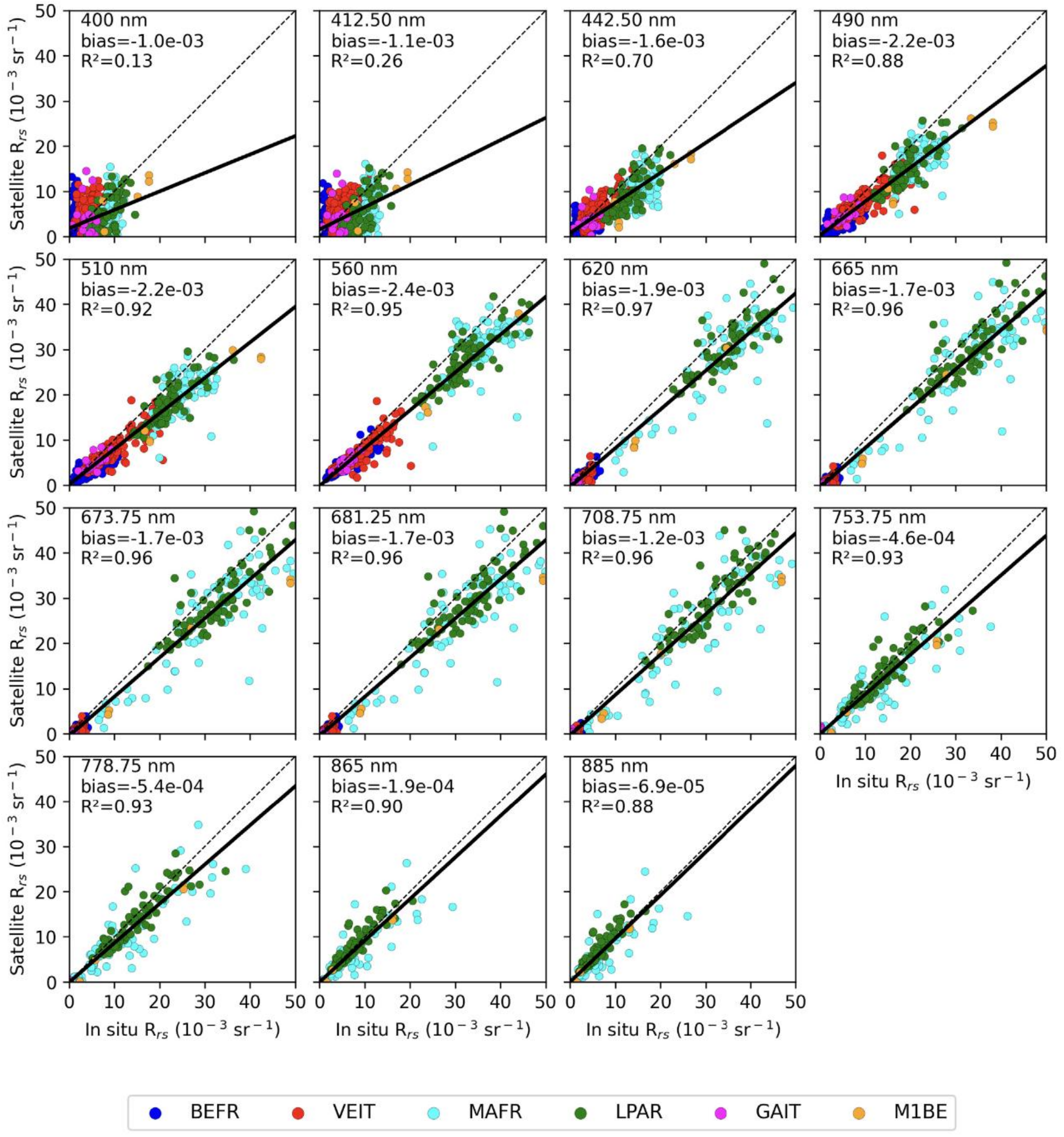


FIGURE 5
Scatter plot of *Rrs* match-ups between satellite (Sentinel-3 WFR) and *in situ* (L2 HYPSTAR®) measurements for each OLCI band. Data points are coloured by site. Statistics are computed including the six sites.

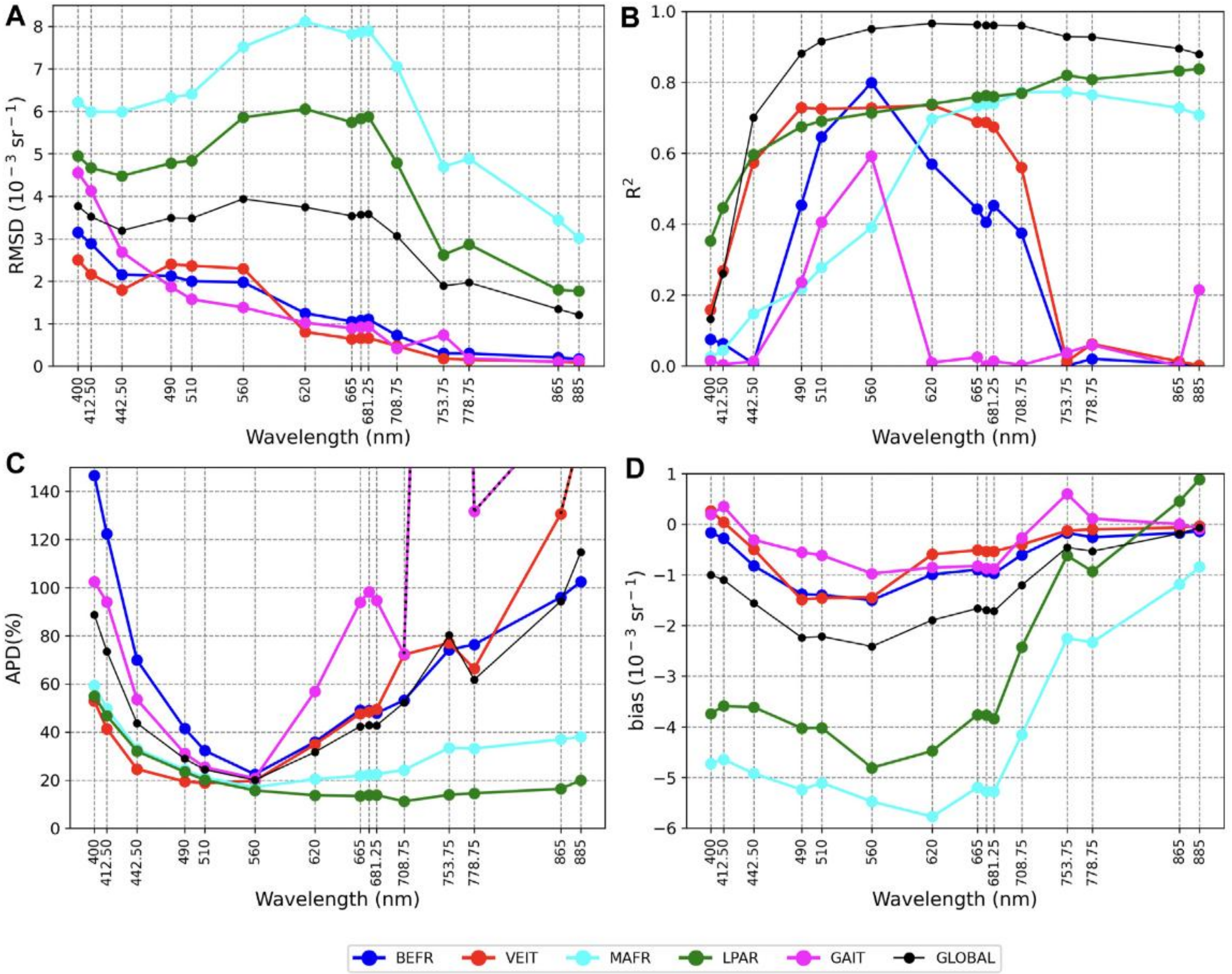


FIGURE 6
Spectral variation of the validation metrics computed for each site from the Sentinel-3 WFR match-ups with HYPSTAR® L2 *in situ* data. **(A)** RMSD (in *Rrs* units: sr^{-1}). **(B)** Determination coefficient (R^2). **(C)** Absolute percentage difference (in percentage). **(D)** Bias (in *Rrs* units: sr^{-1}).

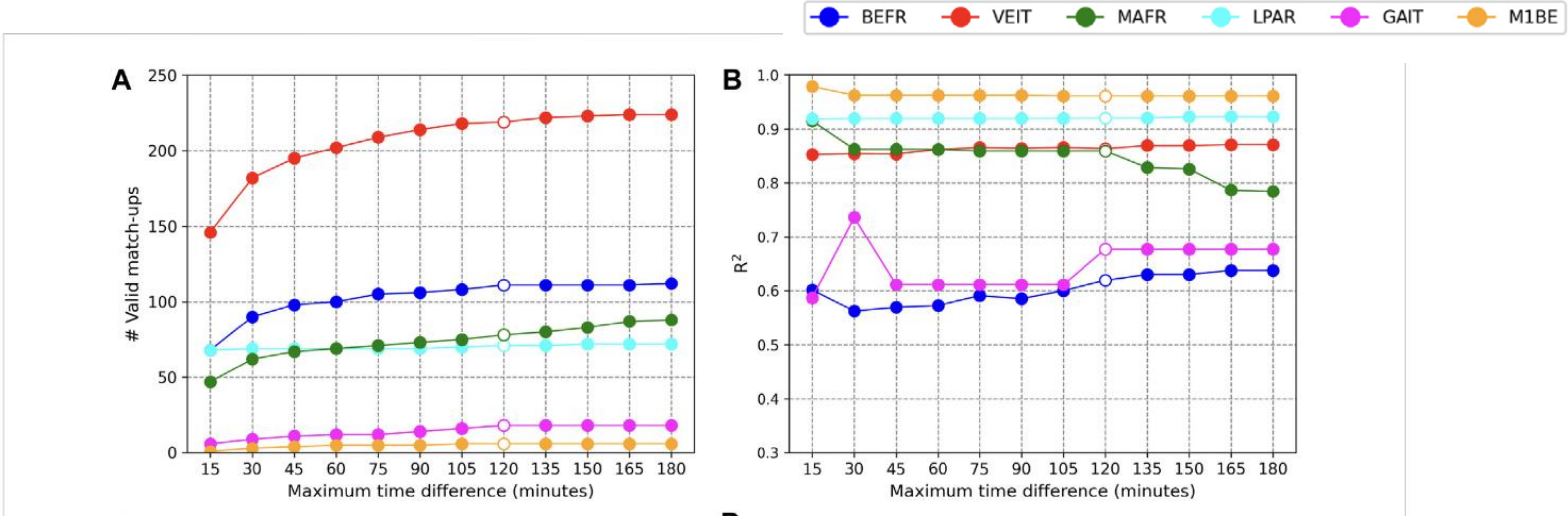
What is the variation of the number of valid match-ups and R^2 using the Sentinel-3 OLCI WFR match-ups with:

1) the maximum time difference between the satellite and in situ measurements

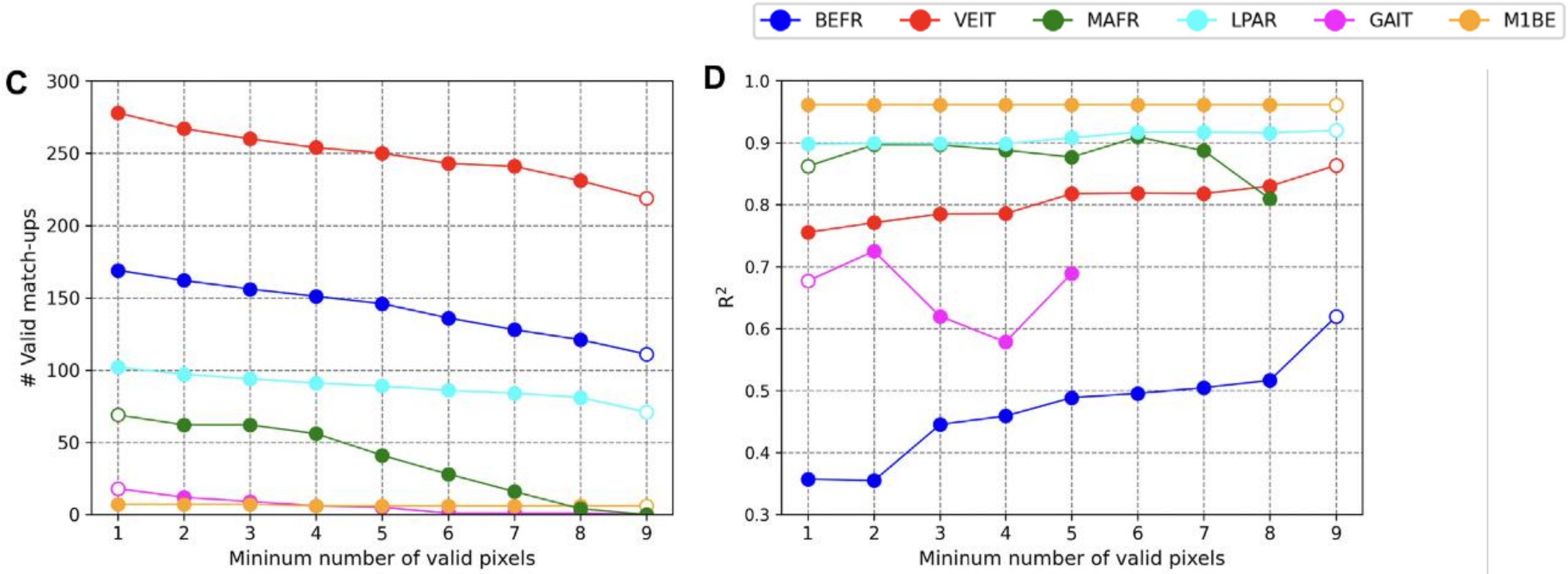
   The use of match-ups with a higher time difference is expected to introduce uncertainties in dynamic environments

2) the minimum number of valid pixels in the satellite extract?

   As the minimum number of valid pixels in the extraction window increases, the number of valid match-ups decreases but validation metrics are expected to improve, since higher uncertainties in the satellite measurement are expected when invalid (masked) pixels are present within the extraction window

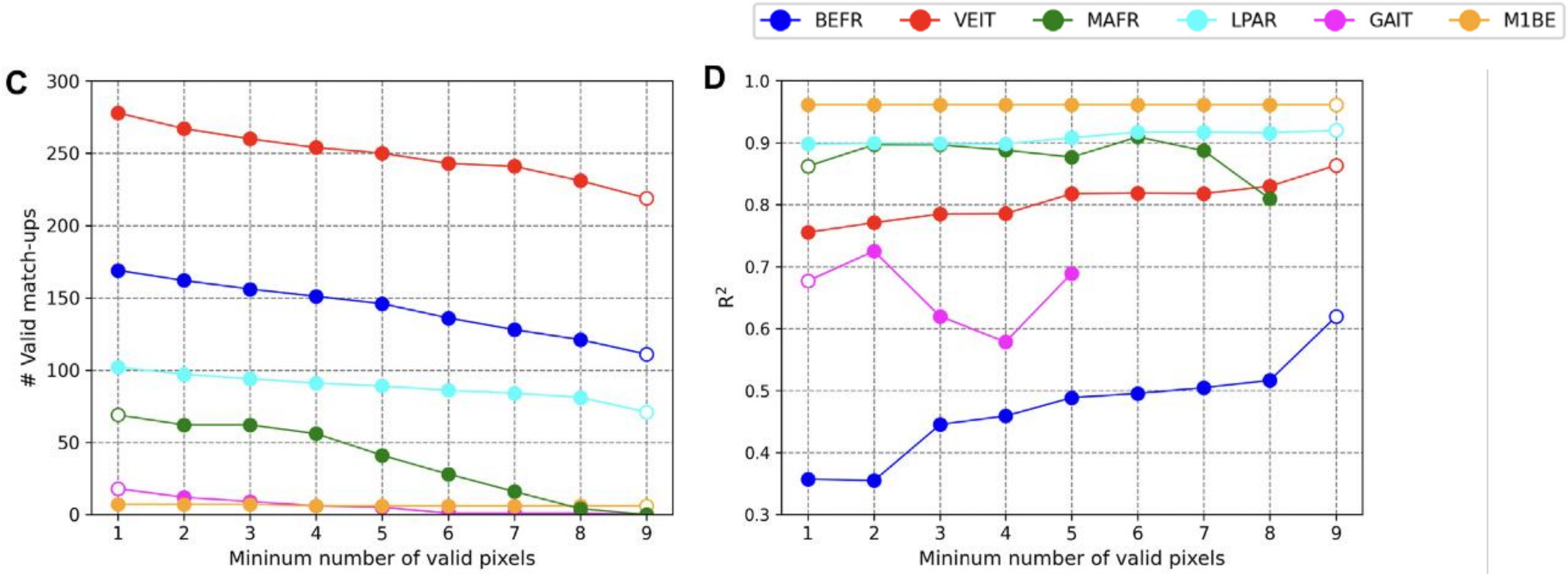


At LPAR, more than 95% of match-ups show a time difference lower than 15 min and R^2 keeps almost constant.
At BEFR, VEIT, and MAFR, around 80% of the valid match-ups were obtained with time differences lower than 30 min with an abrupt change between 15 and 30 min followed by a slower growth.



The strict criterium of 9 valid pixels (i.e., not allowing invalid pixels in the 3 × 3 extraction window) at BEFR, VEIT, LPAR, and M1BE with the **aim of obtaining the best possible validation results at the cost of a lower number of match-ups**

In terms of global correlation, the improvement is more evident at VEIT and mainly at BEFR



the criterium was relaxed at MAFR and GAIT because these two sites are nearer to the coastline so that land masked pixels are always present in the extraction window.

maximizing the number of valid match-ups requiring only one valid pixel

Assessing the influence of different validation protocols on Ocean Colour match-up analyses

- Although the same reference dataset was used, the differences between methods provide a different ”impression of accuracy”.
- The accuracy reported in different studies may not always be directly comparable.
- Details on how the protocols is implemented should be reported every time.

Grazie!

- This work has been performed in the context of the Ocean Colour Thematic Assembly Centre of the Copernicus Marine Environment and Monitoring Service (grant no: 77-CMEMS-TAC-OC), the European Union's Horizon 2020 HYPERNETS Project (grant agreement No 775983), the HYPERNETS-POP project funded by the European Space Agency (contract no 4000139081/22/I-EF), the ArcticFlux TOSCA research project funded by the French Spatial Agency CNES and the ANPCyT PICT-2020/2636 project.
- Thanks to Giuseppe Zibordi for establishing and maintaining the five AERONET-OC sites used in this study.
- Thanks to Kevin Ruddick (RBINS) for his help in the conceptualization of this work.
- Thanks to Simone Colella for helping in the data processing.