Welcome to the
Breakout Workshop on Vicarious Calibration and Validation Protocols

Chairs

Giuseppe Zibordi, Kenneth Voss and B. Carol Johnson

Objective

Reach consensus on standardized protocols for the operational identification and application of in situ measurements to validation and system vicarious calibration (SVC) processes. Consensus should consider the need to apply state-of-the-art methods (e.g., detailed uncertainty budgets for in situ measurements), recognizing practical limitations intrinsic of validation and SVC processes (e.g., the difficulty/impossibility of addressing sub-pixel variability).

The Workshop does not address measurement protocols.

Agenda

- *Introduction to the session on practices for the construction of in situ – satellite matchups, their application to the validation of data products and the presentation of matchup statistics* (G. Zibordi, 15 minutes)
- Group discussion (everybody, tentatively 45 minutes)
-Coffee available (without break)
- *Introduction to the session on practices for the construction of in situ – satellite matchups, their application to SVC and the statistical assessment of derived calibration factors* (K. Voss, 15 minutes)
- Group discussion (everybody, tentatively 45 minutes)
- Wrap-up and consensus consolidation (30 minutes)

Introduction to the breakout discussion on:
practices for the construction of in situ-satellite matchups, their application to the validation of data products and the presentation of matchup statistics

Giuseppe Zibordi, Kenneth J. Voss and B. Carol Johnson

Breakout Workshop on *Vicarious Calibration and Validation Protocols,* 2019 International Ocean Colour Science Meeting,
Busan, South Korea, 9-12 April 2019

Some references on the validation of ocean color data products

Franz, B. A., P. J. Werdell, G. Meister, S. W. Bailey, R. E. Eplee, G. C. Feldman, E. Kwiatkowska, C. R. McClain, F. S. Patt, and D. Thomas (2005), The continuity of ocean color measurements from SeaWiFS to MODIS, in Earth Observing Systems X, Proc. SPIE, vol. 5882, edited by J. J. Butler, Bellingham, Wash.

Mélin, F., Berthon, J. F., & Zibordi, G. (2005). Assessment of apparent and inherent optical properties derived from SeaWiFS with field data. *Remote Sensing of Environment*, *97*(4), 540-553.

Zibordi, G., Mélin, F., & Berthon, J. F. (2006). Comparison of SeaWiFS, MODIS and MERIS radiometric products at a coastal site. *Geophysical Research Letters*, *33*(6).

Mélin, F., Zibordi, G., Berthon, J.-F., 2007. Assessment of satellite ocean color products at a coastal site. Remote Sens. Environ. 110, 192–215. https://doi.org/10.1016/j.rse.2007.02.026.

Antoine, D., d'Ortenzio, F., Hooker, S.B., Bécu, G., Gentili, B., Tailliez, D., Scott, A.J., 2008. Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color sensors (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean Sea (BOUSSOLE project). *J. Geophys. Res.* 113, C07013. https://doi. org/10.1029/2007JC004472.

Hlaing, S., Harmel, T., Gilerson, A., Foster, R., Weidemann, A., Arnone, R., Wang, M., Ahmed, S., 2013. Evaluation of the VIIRS ocean color monitoring performance in coastal regions. Remote Sens. Environ. 139, 398–414. https://doi.org/10.1016/j.rse.2013.08.013.

Wang, M., Liu, X., Jiang, L., Son, S., Sun, J., Shi, W., Tan, L., Naik, P., Mikelsons, K., Wang, X., Lance, V., 2014. Evaluation of VIIRS Ocean Color Products 92610E. https://doi.org/10.1117/12.2069251.

Vandermeulen, R.A., Arnone, R., Ladner, S., Martinolich, P. (2015). Enhanced satellite remote sensing of coastal waters using spatially improved bio-optical products from SNPP-VIIRS. Remote Sens. Environ. 165, 53–63. https://doi.org/10.1016/j.rse.2015.04.026.
Müller, D., Krasemann, H., Brewin, R. J., Brockmann, C., Deschamps, P. Y., Doerffer, R., ... & Mélin, F. (2015). The Ocean Colour Climate Change Initiative: II. Spatial and temporal homogeneity of satellite data retrieval due to systematic effects in atmospheric correction processors. *Remote Sensing of Environment*, *162*, 257-270

Moore, T. S., Campbell, J. W., & Feng, H. (2015). Characterizing the uncertainties in spectral remote sensing reflectance for SeaWiFS and MODIS-Aqua based on global in situ matchup data sets. *Remote Sensing of Environment*, *159*, 14-27.

Brando, V.E., Lovell, J.L., King, E.A., Boadle, D., Scott, R., Schroeder, T., 2016. The potential of autonomous ship-borne hyperspectral radiometers for the validation of ocean color radiometry data. Remote Sens. 8. https://doi.org/10.3390/rs8020150.

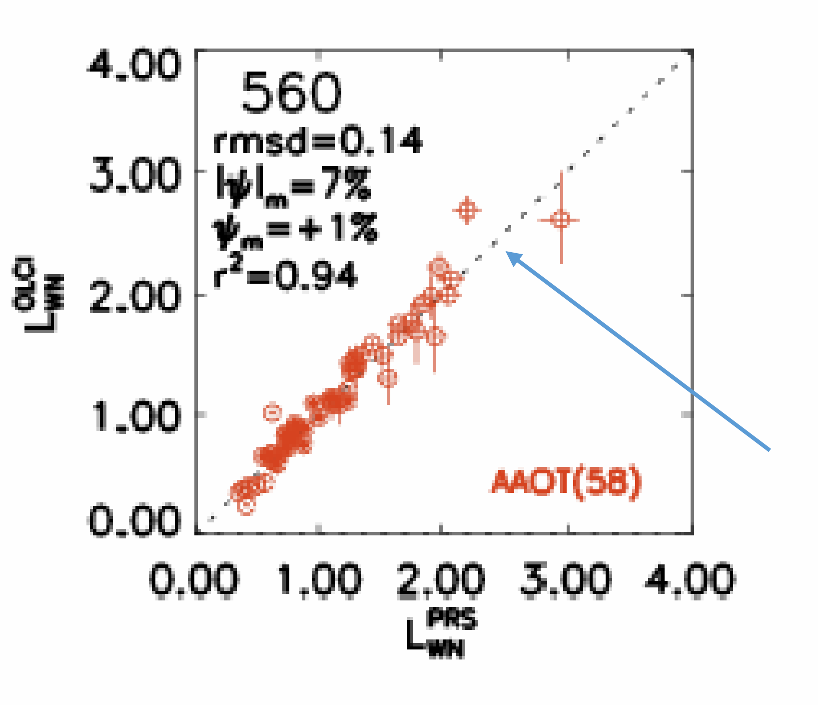
Gerbi, G. P., Boss, E., Werdell, P. J., Proctor, C. W., Haëntjens, N., Lewis, M. R., ... & Koegler, J. (2016). Validation of ocean color remote sensing reflectance using autonomous floats. *Journal of Atmospheric and Oceanic Technology*, *33*(11), 2331-2352.

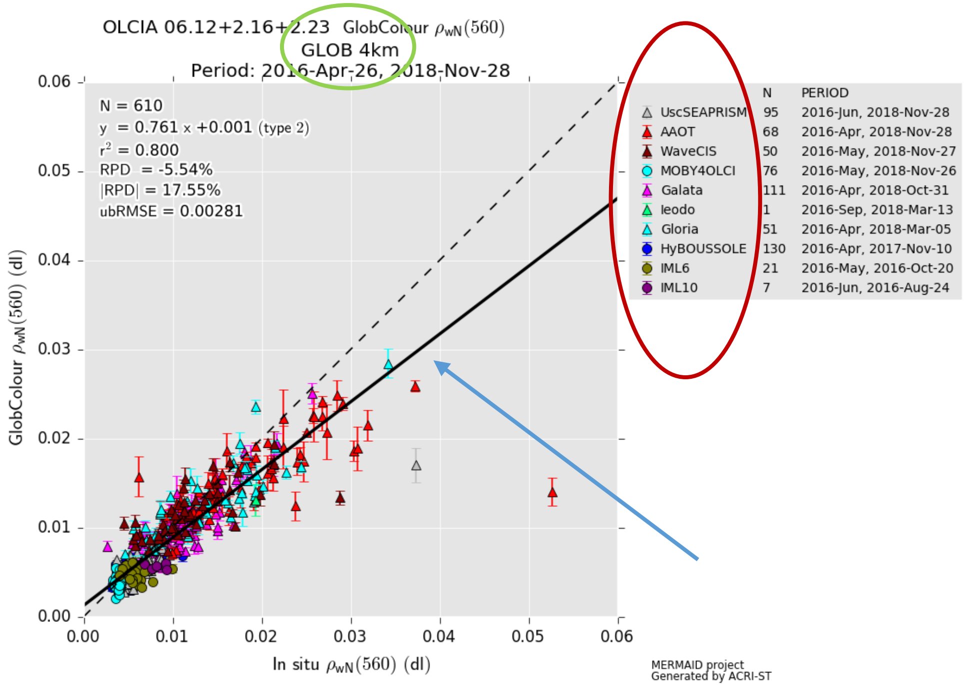
Pahlevan, N., Schott, J. R., Franz, B. A., Zibordi, G., Markham, B., Bailey, S., ... & Strait, C. M. (2017). Landsat 8 remote sensing reflectance (Rrs) products: Evaluations, intercomparisons, and enhancements. *Remote Sensing of Environment*, *190*, 289-301.

Seegers, B. N., Stumpf, R. P., Schaeffer, B. A., Loftin, K. A., & Werdell, P. J. (2018). Performance metrics for the assessment of satellite data products: an ocean color case study. *Optics express*, *26*(6), 7404-7422.

Zibordi, G., Mélin, F., & Berthon, J. F. (2018). A Regional Assessment of OLCI Data Products. *IEEE Geoscience and Remote Sensing Letters*, (99), 1-5.

On the need of standardizing validation (1)

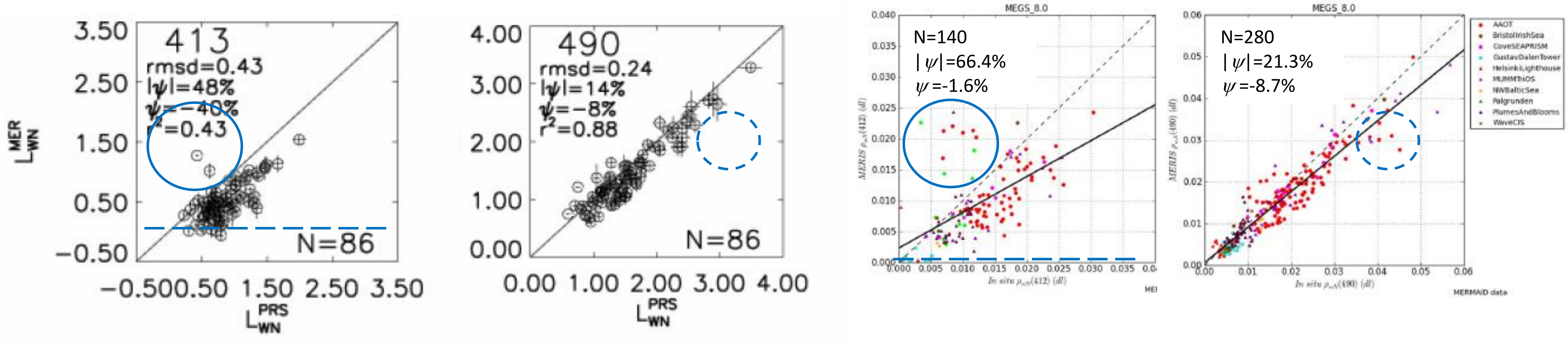




Validation data from the same site and data source appear to lead to a different trend!

Are the validation sites applicable for 4km satellite data? Is the distance from the coast properly considered?

On the need of standardizing validation (2)

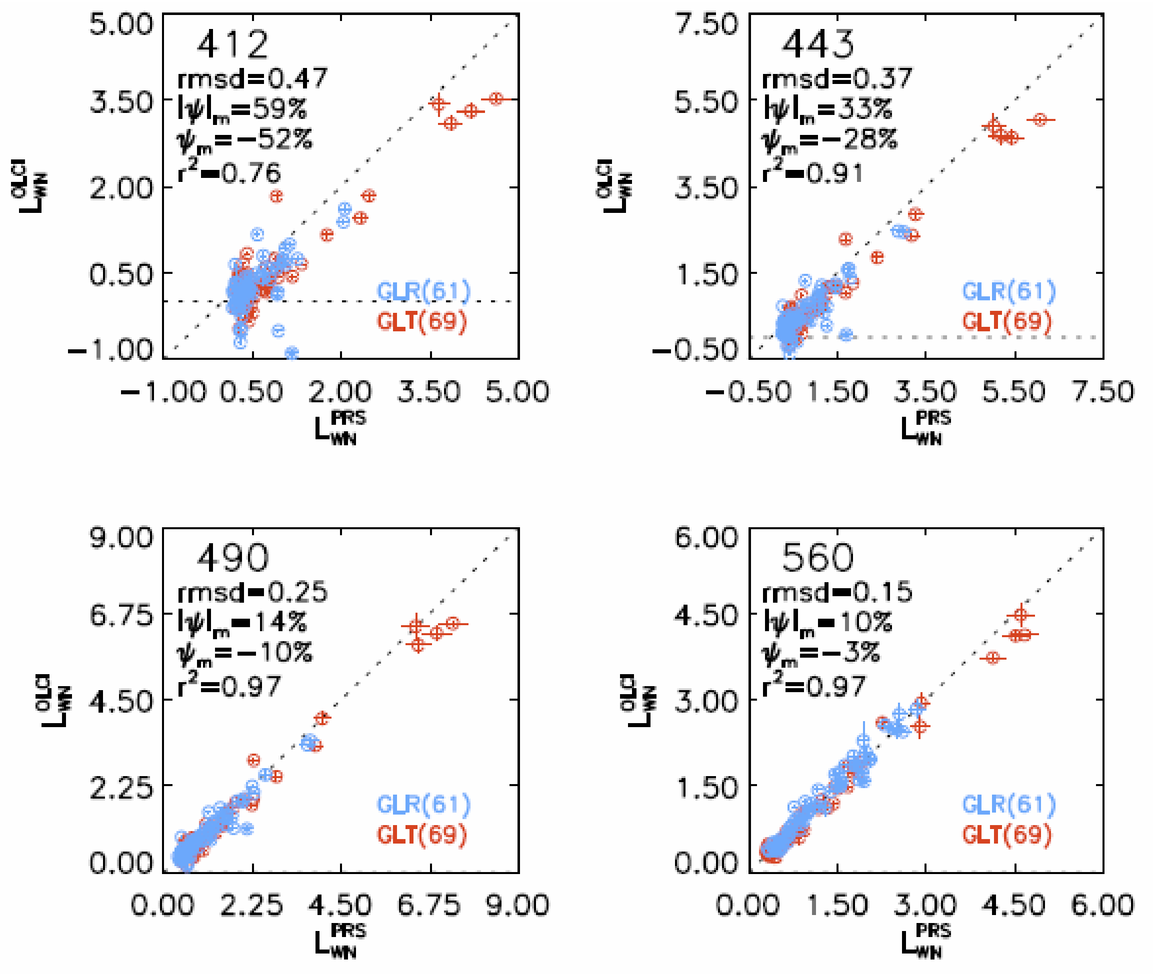


Focusing on the AAOT data (that exhibits the largest range of values), which are common to both assessments, note the following differences:

1. The exclusion of negative values in the scatter plots for 412 nm displayed on the right;
2. The clear application of different quality criteria evident for high values.

Validation results become quite spectrally different!

On the need of standardizing validation (3)



Flags are commonly mission (source) specific.

The example shows the impact (highlighted by empty circles) of the application of a specific flag to OLCI validation data. Apparently the flag tends to exclude data exhibiting high radiance values. Thus, the application or non application of the flag may lead to validation results that, aside potentially leading to different statistical results, would refer to different ranges of input radiances.

The application of flags cannot be subjective, but flags need to be fully justified.

Objective of the discussion

- Agree on fundamental requirements for in situ measurements supporting single missions for regional/global applications or multiple-missions addressing climate studies (e.g., geophysical quantities, spectral characteristics, uncertainty budget and traceability, geographical relevance, …).

- Agree on fundamental methods to enforce equivalence of satellite and in situ data (e.g., application of identical corrections for BRDF effects, corrections for minimizing the impact of different spectral bands, …).

- Agree *i.* on criteria for the construction of matchups (e.g., local spatial/ temporal variability, observation conditions, time-lags between in situ and satellite data…) and additionally, *ii.* on methods and criteria for the statistical analysis of matchup data and the following presentation of summary results (e.g., the statistical methods for the determination of systematic differences and dispersions affecting satellite data with respect to in situ measurements …).

Talking points

Talking points ***(separated for Regional/Environmental and Global/Climate applications)***	
In Situ Data: Generic	Quantities, measurement methodology, illumination conditions, distance from the coast, bottom depth, water type, multiple sites/sources
In Situ Data: Radiometric	Uncertainties, spectral resolution, spectral matching, BRDF corrections, band-shift corrections
Matchups Construction	Number of image elements, time-lag, agency suggested flags, viewing and illumination geometries, thresholds on variation coefficients
Matchup Statistics	Minimum number of matchups, bias index, dispersion index, rmsd, ranges, distributions

In Situ Data: Generic

Talking Point	Applications (generalized)		Relevant references	Notes/Comments
In Situ Data: Generic	***Regional /Environmental***	***Global/Climate***		
Quantity	Radiometry (e.g., Rrs, Lwn) & derived (e.g., Chla, t_a)	Radiometry (e.g., Rrs, Lwn) & derived (e.g., Chla, t_a)		Chla and any derived quantity should be included in the validation process only if the related radiometric product is qualified for validation
Measurement methodology	Declared and documented	Consolidated (sharing community consensous on measurement protocol, data QA/QC and processing)		
Illumination conditions	Clear sky (clouds away from the sun and coverage ideally not exceeding 2-octas)	Clear sky (clouds away from the sun and coverage ideally not exceeding 2-octas)		
Distance from the coast	Declared	Avoiding cases affected by adjacency effects (distance from land should be larger than at least 5 nautical miles)	*Bulgarelli and Zibordi, 2018. JRC Technical Report, doi: 10.2760/178467 (online), 40 pp.*	
Bottom depth	Declared	Avoid cases affected by bottom effects (which depend on depth and water type)		
Water type	Any	Prioritizing mesotrophic/oligotrophic (but not excluding different water type assuming a statistical balance in their represenattivity)		
Multiple sites/sources	Yes	Yes		

International Ocean Colour Science Meeting 2019 | Advancing Global Ocean Colour Observations

In Situ Data: Radiometric

Talking Point	Applications (generalized)		Relevant references	Notes/Comments
In Situ Data: Radiometric	***Regional /Environmental***	***Global/Climate***		
Uncertainties	Declared and documented	Fulfilling GCOS requirements for Rrs and Lwn (i.e., lower than 5% for Rrs in the blue-green spectral regions), lower than 0.02 for t_a, and than ~15% for Chla in the 0.01-10 μg l^{-1} range.	*WMO, 2011. Report GCOS 154, 138 pp.*	
Spectral resolution	Comparable to that of the space sensor	At least comparable to that of the space sensor (typically 10 nm or better)		
Spectral matching	Desirable to within a few nm	Required (i.e., in situ and satellite equivalent center-wavelengths closer than 2-5 nm, depending on the spectral location of the band)		
BRDF corrections	Required (implying corrections equivalent to those applied to satellite data)	Required (implying corrections equivalent to those applied to satellite data)		Still recognizing that corrections non suitable for the specific water type may become an additional source of uncertainty
Band-shift corrections	Desirable in the full visible spectrum for center-wavelengths differing by more than 1-2 nm	Required in the full visible spectrum for center-wavelengths differing by more than 1-2 nm (implying direct or indirect knowledge of local IOPs)	*Melin and Sclept, 2015. Optics Express, 23, 2262-2279.*	

Matchups Construction

Talking Point	Applications (generalized)		Relevant references	Notes/Comments
Matchups Construction	***Regional /Environmental***	***Global/Climate***		
Number N of image elements centered at the validation site (1-element ~ 1km for reduced resolution data)	Tentatively 3x3 in coastal (in view of accounting for coastal variability and minimize land perturbations) and 5x5 in open sea regions	Tentatively 3x3 in coastal (in view of accounting for coastal variability and minimize land perturbations) and 5x5 in open sea regions		In the case of high spatial resolution satellite data, it is difficult to propose general values of N suppoted by published work (there are cases considering 3x3 and others juts 1)
Time-lag between satellite and in situ data	Less than 4 hr (still, most suitable value should be determined accounting for local variability)	Less than 2 hr (sensitivity tests based on different time-lags, may provide elements in support of the selected value)		
Agency Suggested Flags	All (each one not affecting any of the N image elements: i.e., 100% of the N elements should not be affected by suggested flags)	All (each one not affecting any of the N image elements: i.e., 100% of the N elements should not be affected by suggested flags)		
Viewing and illumination geometries	Viewing angle and sun zenith lower than given threshods (e.g., 60 and 70 degrees)	Viewing angle and sun zenith lower than given threshods (e.g., 60 and 70 degrees)		
Threshold on the variation coefficient of the N elements	Tentatively 0.2 at a single spectral band (e.g., 555 nm or equivalent for Rrs or Lwn, and 870 nm or equivalent for t_a)	Tentatively 0.2 at a single spectral band (e.g., 555 nm or equivalent for Rrs or Lwn, and 870 nm or equivalent for ta). Sensitivity tests may support the selected value.		Thresholds on the variation coefficient of in situ data over periods of n*time-lags, may additionally help removing cases affected by high temporal (spatial) variability

Matchup statistics

Talking Point	Applications (generalized)		Relevant references	Notes/Comments
Matchups Statistics	***Regional /Environmental***	***Global/Climate***		
Minimum number of matchups (for a given processing and period)	No (but still enough to assume statistical representativity of regional spatial/temporal variability)	Ensuring statistical representativity (tentatively more than several hundreds)		When satisfying statistical representativity, matchups should be constructed and analyzed for different cases
Bias index (for each band and all matchups)	Computed from the same matchups for all visible bands or products (e.g., median of percent differences)	Computed from the same matchups for all visible bands or products (e.g., median of percent differences)		
Dispersion index (for each band and all matchups)	Computed from the same matchups for all visible bands or products (e.g., median of percent absolute differences)	Computed from the same matchups for all visible bands or products (e.g., median of percent absolute differences)		
Root mean square of differences (for each band and all matchups)	Desirable	Computed from the same matchups for all visible bands or products		
Ranges	Required (essential to determine the comparability of results across indedendent analysis from different geographic regions and water types)	Required (essential to determine the comparability of results across indedendent analysis from different geographic regions and water types)		
Distributions	Desirable	Required for all visible bands or products (essential to determine the significance of statistics)		