# Overcoming Dimensionality Barriers in Ocean Color Data


**Ryan Vandermeulen**

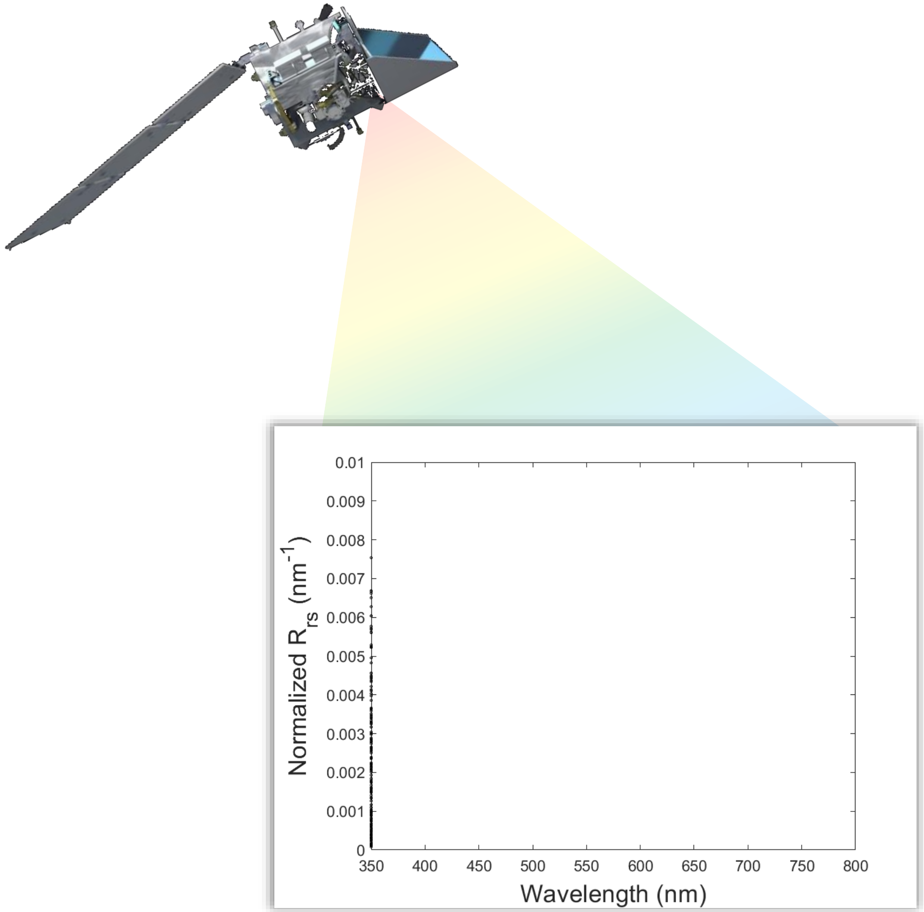
NMFS Satellite Coordinator

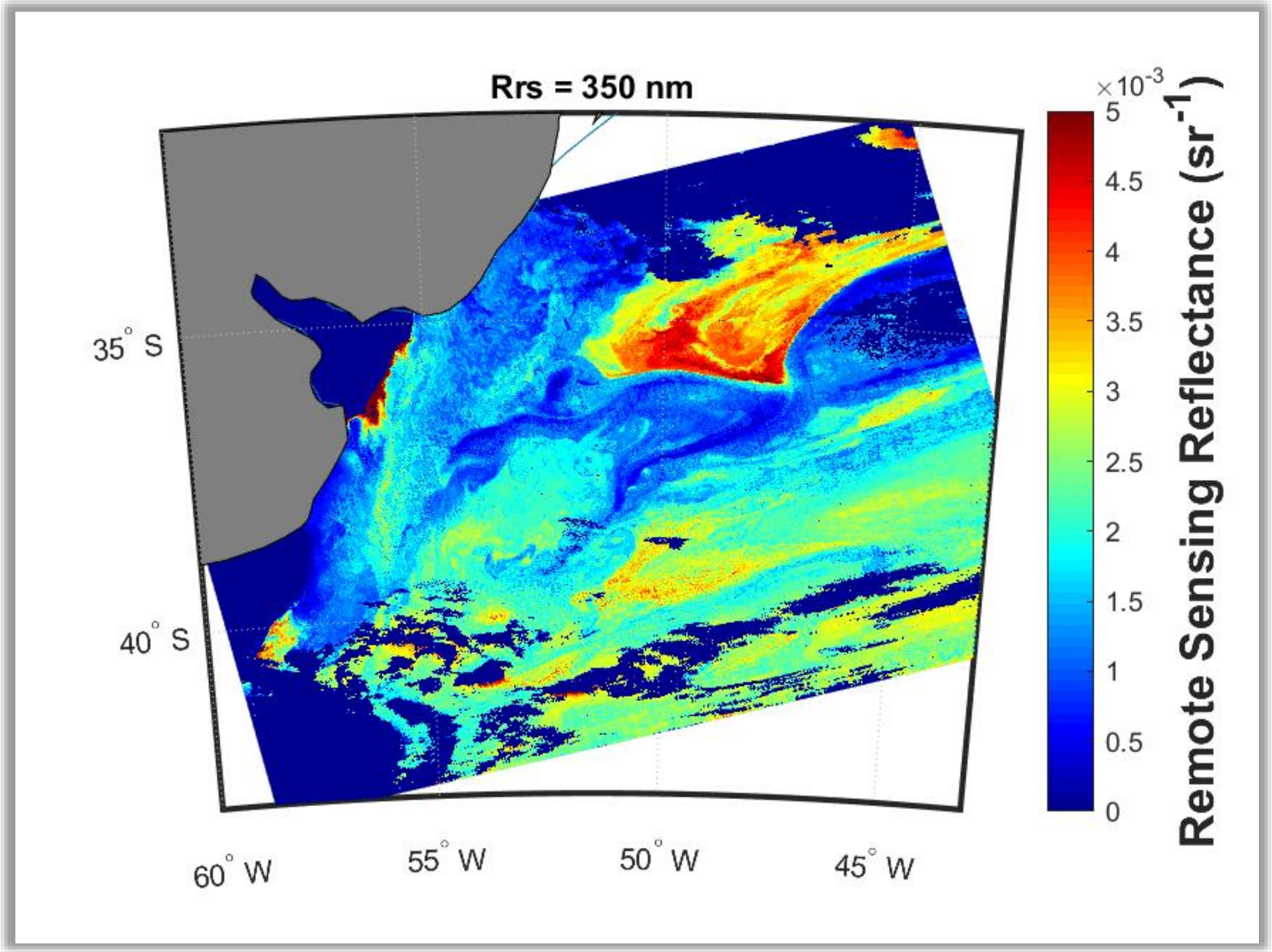
Office of Science & Technology

NOAA National Marine Fisheries Service


**NOCCG Seminar – April 26, 2023**

How can we better contextualize the differences between multiple layers of spectral information with a quantitative metric?




*Example of dynamic range of hyperspectral data for optically complex waters*

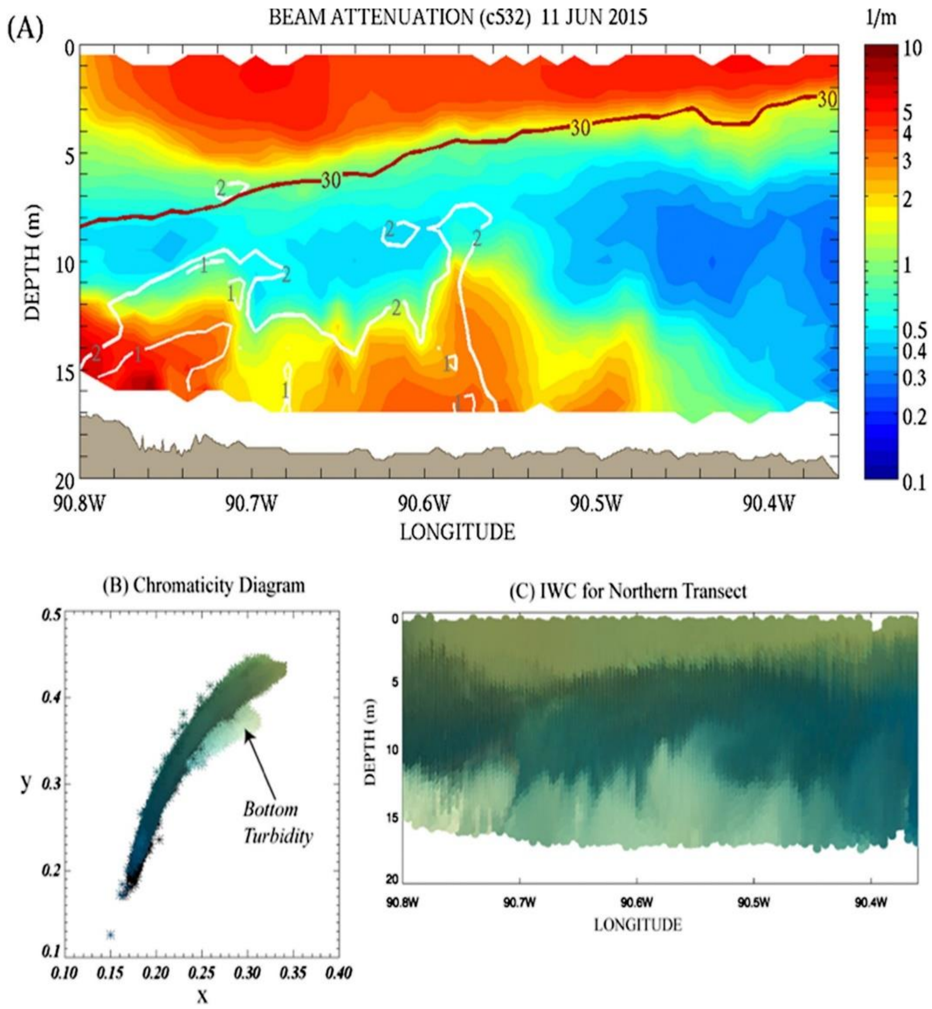
# Spectral Classification (abridged history)

- **Fuzzy c-means classification** (Moore et al. 2009, 2014, Eleveld et al. 2017)
- **Agglomerative Ward's linkage** (Lubac and Loisel 2007)
- **Iterative self-organizing data analysis technique** (Melin and Vantrepotte 2015)
- **Varimax-rotated Principal Component Analysis** (Avouris and Ortiz 2019)
- **Max-classification** (Ye et al. 2016)
- **K-means clustering** (Wei et al. 2022, Prasad and Agarwal 2016)
- **Chromaticity coordinates/Hue angle CIE** (Wernand et al. 2013, Jolliff et al. 2018)

## POTENTIAL AREAS OF OPPORTUNITY

- Usually require training data sets (hard to come by)
- Often yield regionally-specific results, requires development
- Arbitrarily define number of clusters in some instances
- Data output may be dimensionless (Class A, B, C, etc.)
- Challenging to interpret without some a priori knowledge of the data
- May not be consistent across various sensor configurations
- Incremental output creates sharp boundary layers between water types

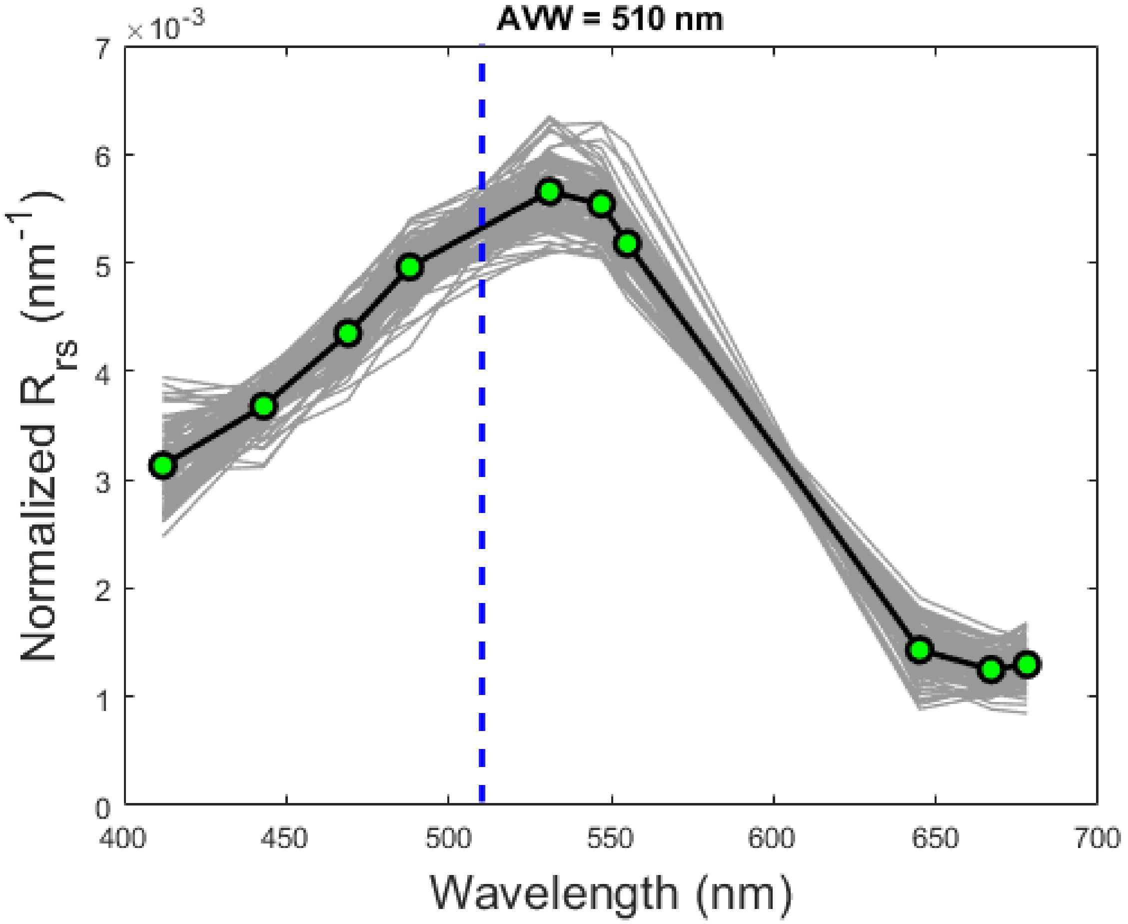
***Jolliff et al. 2018***



# Apparent Visible Wavelength

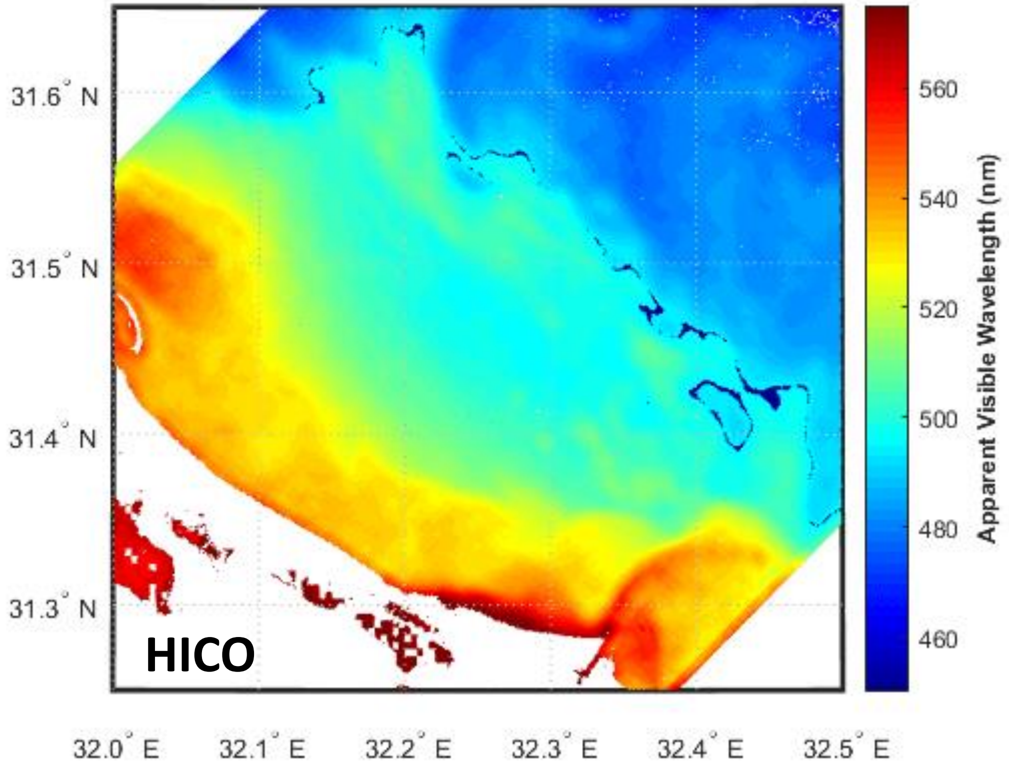
$$AVW = \frac{\sum_{i=\lambda_1}^{\lambda_n} R_{RS}(\lambda_i)}{\sum_{i=\lambda_1}^{\lambda_n} \frac{R_{RS}(\lambda_i)}{\lambda_i}} = \left( \frac{\sum_{i=\lambda_1}^{\lambda_n} \lambda_i^{-1} R_{RS}(\lambda_i)}{\sum_{i=\lambda_1}^{\lambda_n} R_{RS}(\lambda_i)} \right)^{-1}$$

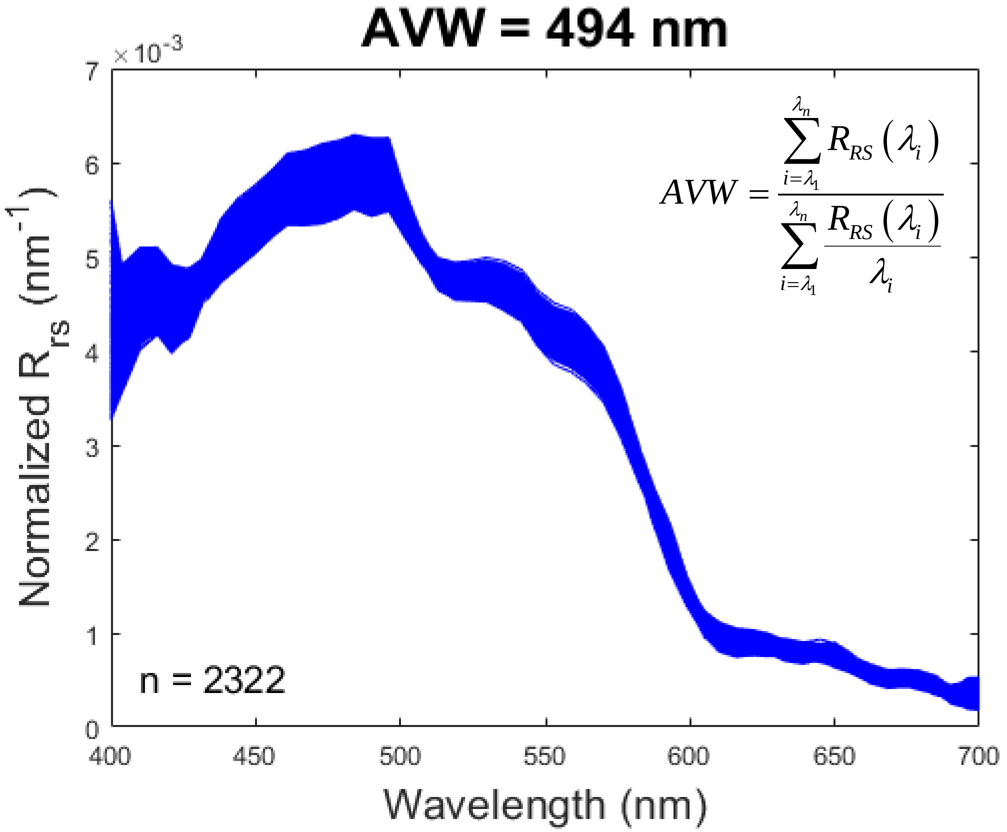
*The weighted harmonic mean of the $R_{RS}$ wavelengths, outputs an* ***Apparent Visible Wavelength, AVW*** *(in units of nm). The derivation of the AVW is simply a first-order measure of the dominant color of the water, as determined by the weight that each measured channel contributes to the reflectance in the visible range of the spectrum.*




*Vandermeulen et al. (2020). 150 Shades of Green: Using the full spectrum of remote sensing reflectance to elucidate color shifts in the ocean. Remote Sensing of Environment.* [https://doi.org/10.1016/j.rse.2020.111900](https://doi.org/10.1016/j.rse.2020.111900)

# New tools to investigate full spectral behavior

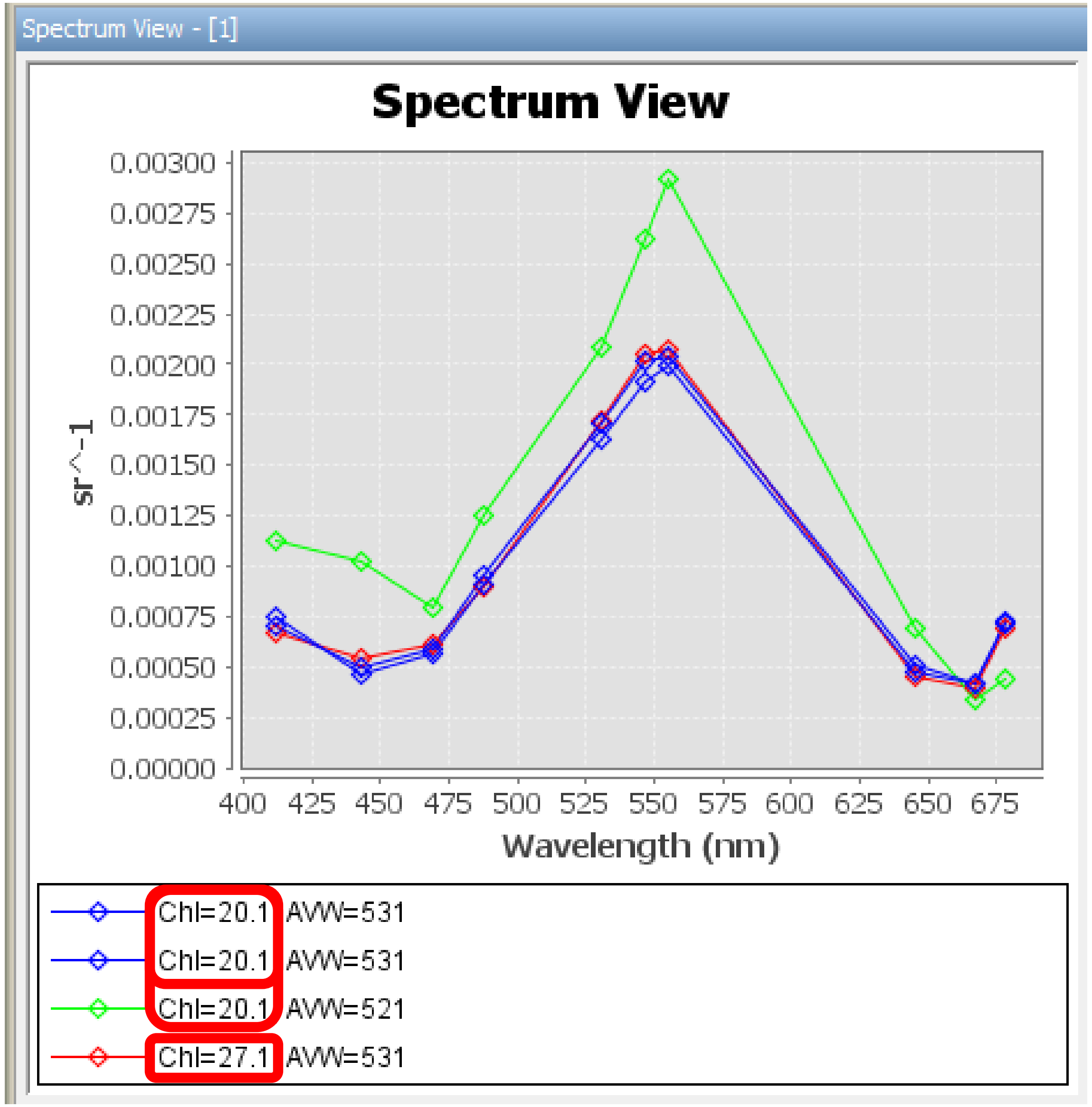


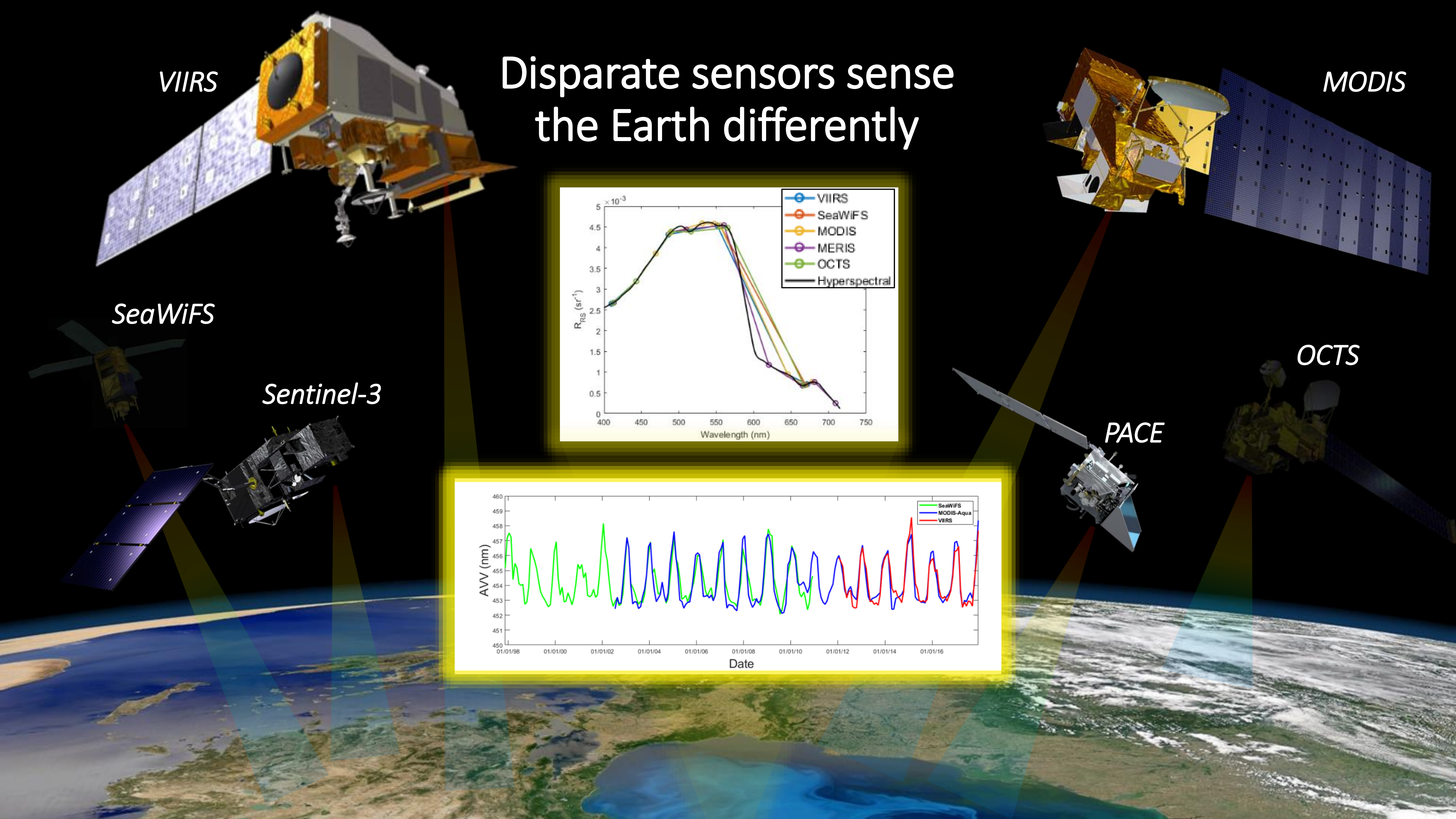

*Using full spectral information represents a more holistic approach to unraveling spectral variability, ensuring that any diagnostic signals present are considered, and thus can help maximize the potential of spectral information embedded in remote sensing data.*

# Is this really that different from other OC products?

- AVW is the only algorithm to utilize the entire visible spectrum of $R_{rs}$, and thus is a useful diagnostic tool for assessing spectral deviations.
- Two **identical** chl-a values can exhibit vastly different spectral shapes
  - green v. blue lines
- Two very **different** chl-a values can exhibit very similar spectral shapes
  - red v. blue lines
- Spectral shape and magnitude metrics can inform specific information about the *behavior* of $R_{rs}$.



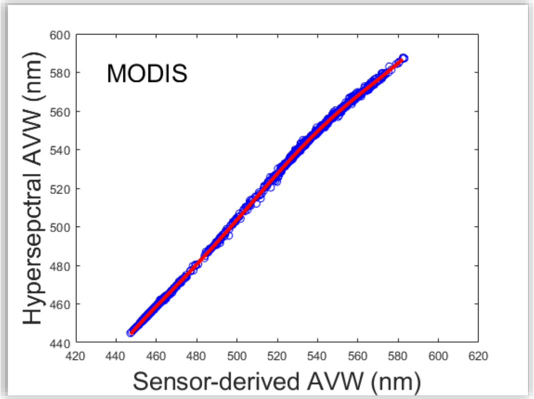
# Disparate sensors sense the Earth differently



# Sensor-specific *AVW* calibration

| | |
|---|---|
| MODIS-Aqua | 412, 443, 469, 488, 531, 547, 555, 645, 667, 678 |
| MODIS-Terra | 412, 443, 469, 488, 531, 547, 555, 645, 667, 678 |
| OLCI-S3A | 400, 412, 443, 490, 510, 560, 620, 665, 674, 682 |
| OLCI-S3B | 400, 412, 443, 490, 510, 560, 620, 665, 674, 681 |
| MERIS | 413, 443, 490, 510, 560, 620, 665, 681 |
| SeaWiFS | 412, 443, 490, 510, 555, 670 |
| HawkEye | 412, 447, 488, 510, 556, 670 |
| OCTS | 412, 443, 490, 516, 565, 667 |
| GOCI | 412, 443, 490, 555, 660, 680 |
| VIIRS-SNPP | 410, 443, 486, 551, 671 |
| VIIRS-JPSS1 | 411, 445, 489, 556, 667 |
| CZCS | 443, 520, 550, 670 |
| MSI-S2A | 443, 490, 560, 665 |
| MSI-S2B | 443, 490, 559, 665 |
| OLI | 443, 482, 561, 655 |

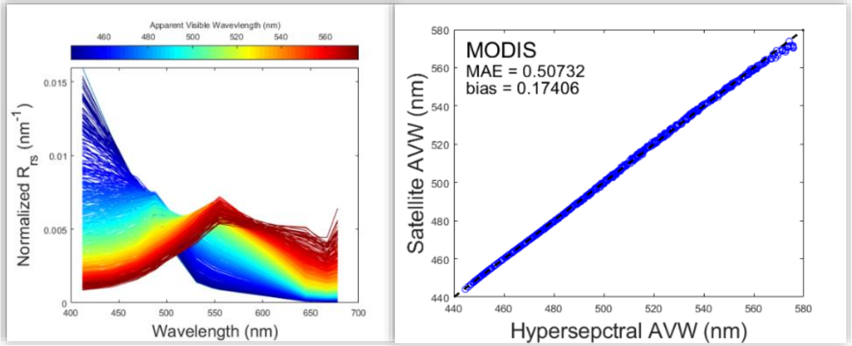
Using defined spectral response functions for each sensor, $R_{rs}(\lambda)$ for each multispectral sensor are reconstructed from the 5,522 in situ hyperspectral spectra. The hyperspectral AVW is compared to the multispectral AVW calculated for each sensor, and the coefficients from a polynomial fit are retained and applied to multispectral satellite data.




| Polynomial Coefficients | | | | | | |
|---|---|---|---|---|---|---|
| Sensor | c0 | c1 | c2 | c3 | c4 | c5 |
| MODIS-Aqua | 5.3223151E-09 | -1.3619239E-05 | 1.3886726E-02 | -7.0534823E+00 | 1.7860303E+03 | -1.8010144E+05 |

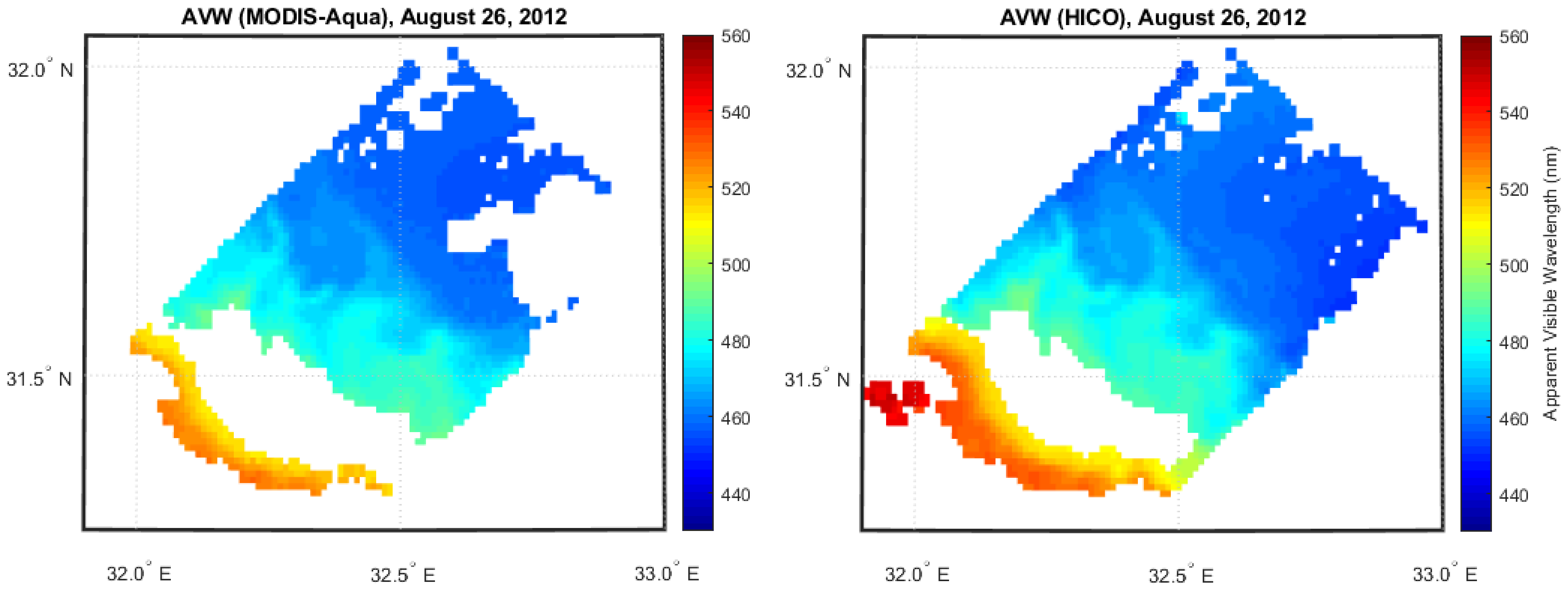
$$AVW_{calibrated}=c_0(AVW^5)+c_1(AVW^4)+c_2(AVW^3)+c_3(AVW^2)+c_4(AVW)+c_5$$

The coefficients are then tested against an independent (synthetic) hyperspectral dataset. The AVW derived from multispectral sensors should now align with a hyperspectral-equivalent value.



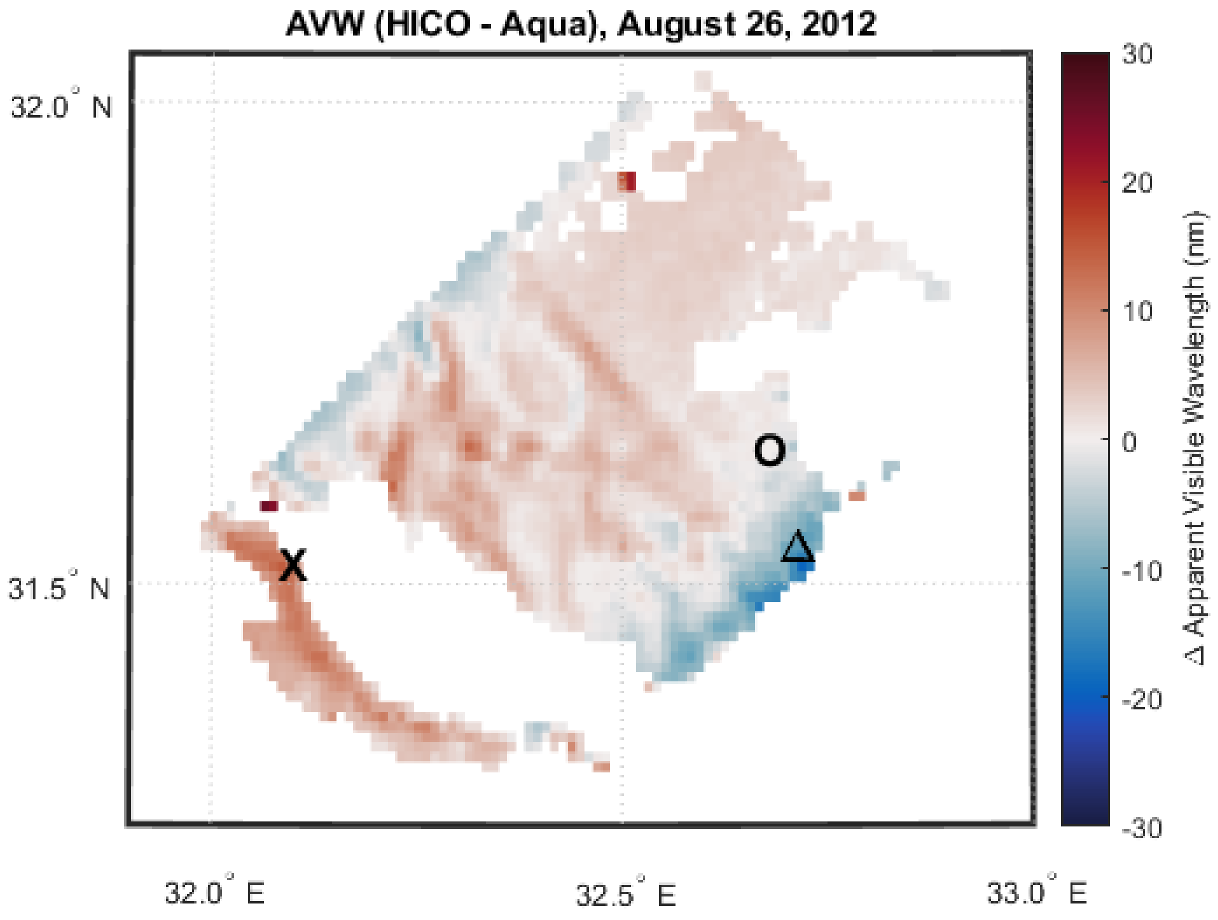

https://oceancolor.gsfc.nasa.gov/atbd/avw/

# Using the Full Spectrum of $R_{rs}(\lambda)$ to Elucidate Color Shifts in the Ocean

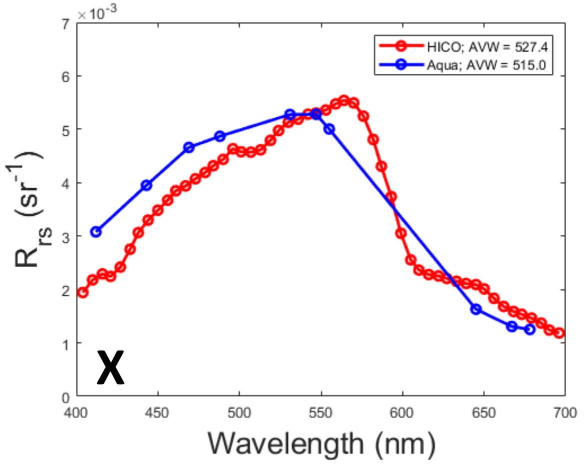



*Apparent Visible Wavelength enables a direct line of comparison for spectral information, fostering synergistic analyses of spectral behavior across satellite and in situ platforms.*

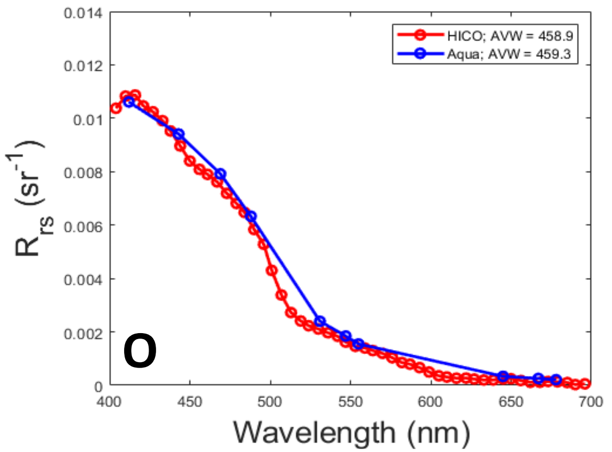
*The dimensionality reduction of spectral information can be used to easily visualize the **directionality** and **magnitude** of spectral shifts over time/space.*



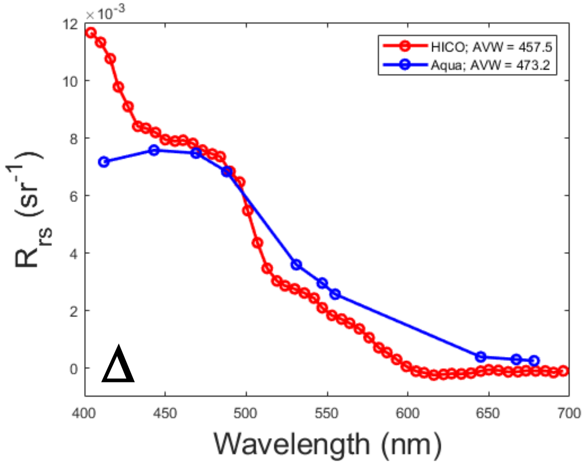

***HICO is red-shifted***




***Spectral Match***




***HICO is blue-shifted***

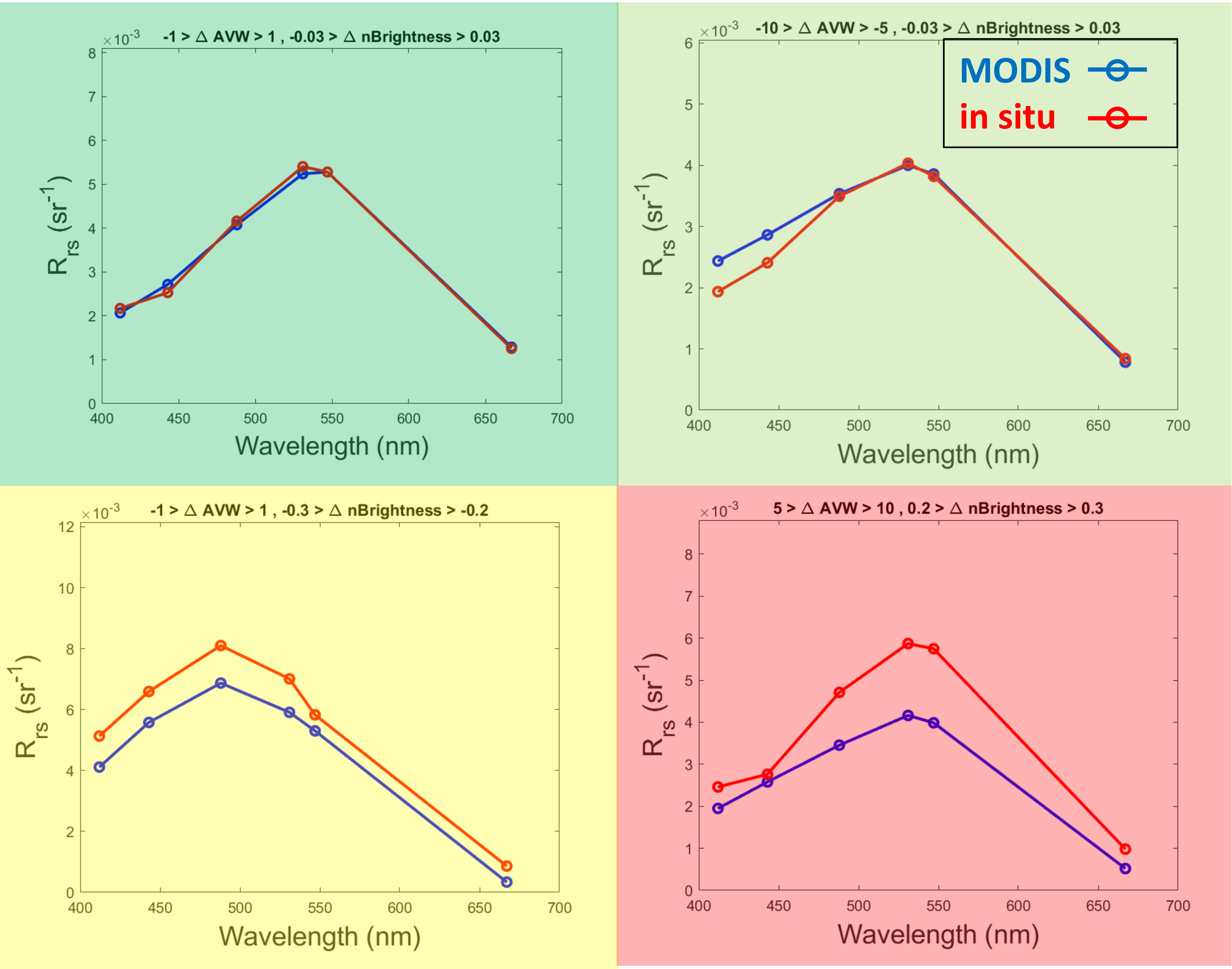


| | |
|---|---|
| CLASS_1 | CLASS_2 |
| CLASS_3 | CLASS_4 |

Initial proof of concept: Downloaded AERONET + *in situ* $R_{rs}$ matchups from SeaBASS, *n = 5,275*

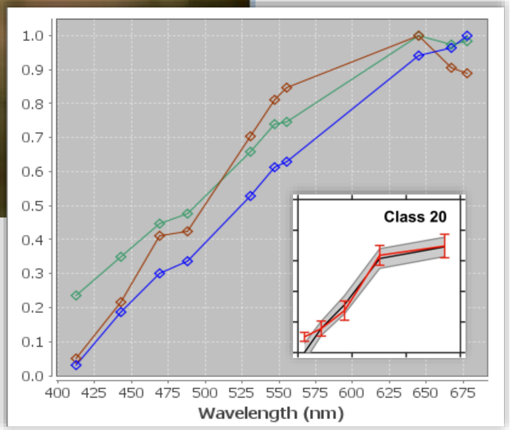
Using full spectral information as a diagnostic tool, we can more quantitatively separate spectra according to the matchup characteristics. Are the spectra overall similar, but the magnitude is off? Are we getting a good "matchup" but detecting seemingly different water types? **NOTE:** not all channels are bad, when we get Class 4 matches, but this may let us categorize statistical analyses for better diagnostics.

**Shape & Magnitude align** | **Shapes different, same magnitudes**

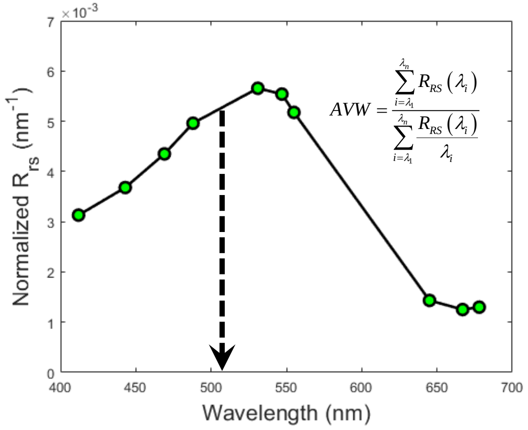


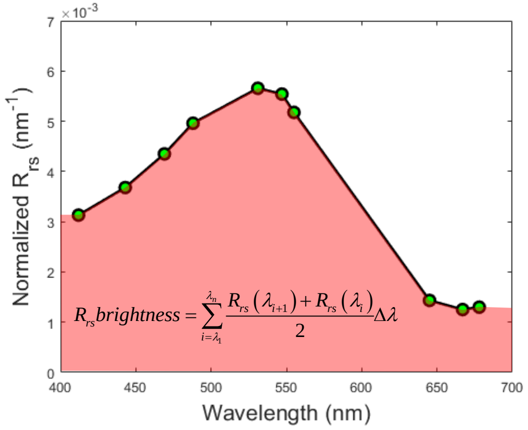

**Shapes same, different magnitudes** | **Shapes different, magnitudes different**

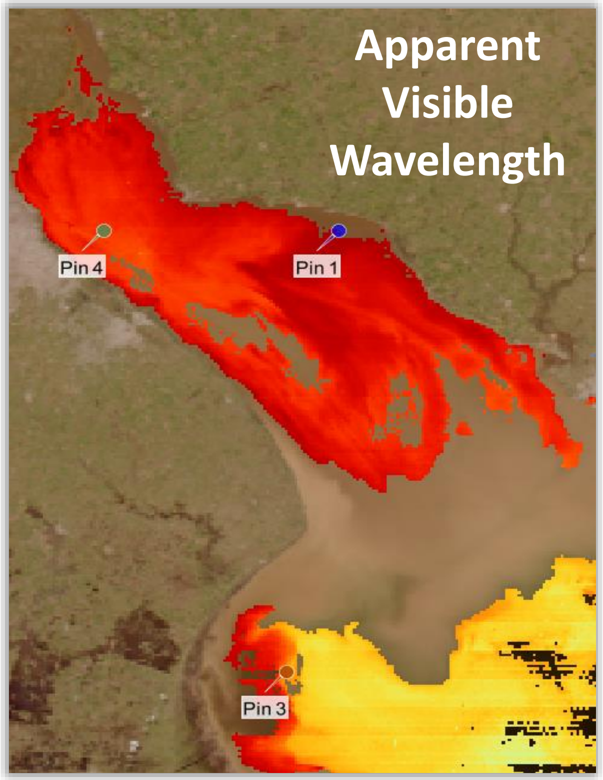
## Using more spectral information elucidates more trends

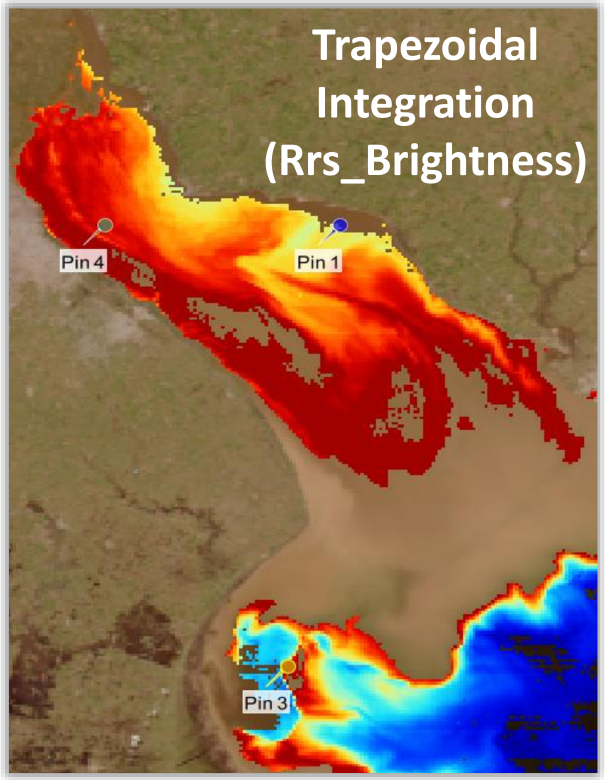




*OWT classification based on spectral shape alone can be limiting, as shown for the Rio de la Plata.*

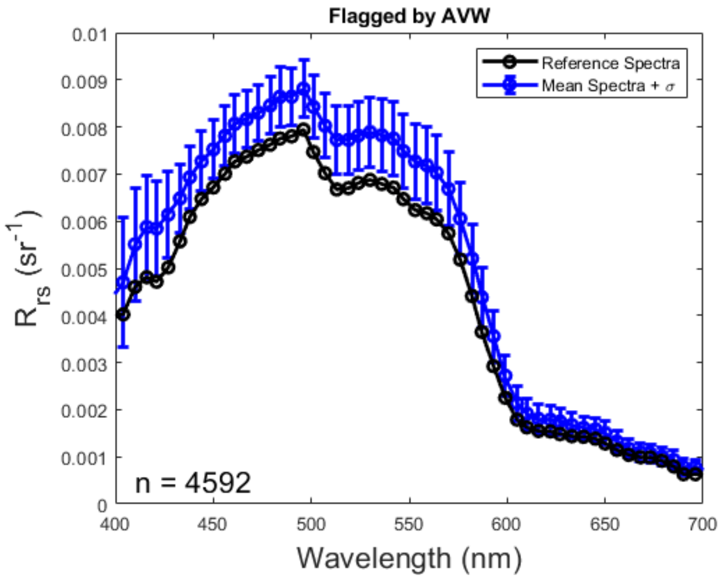


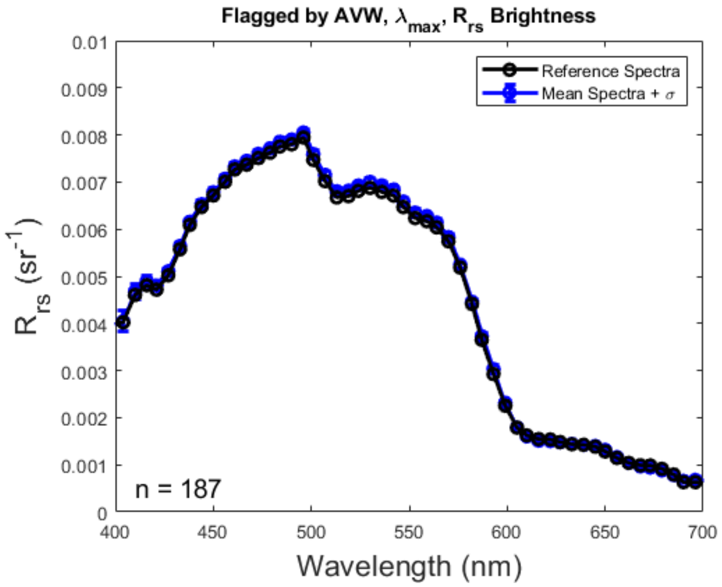



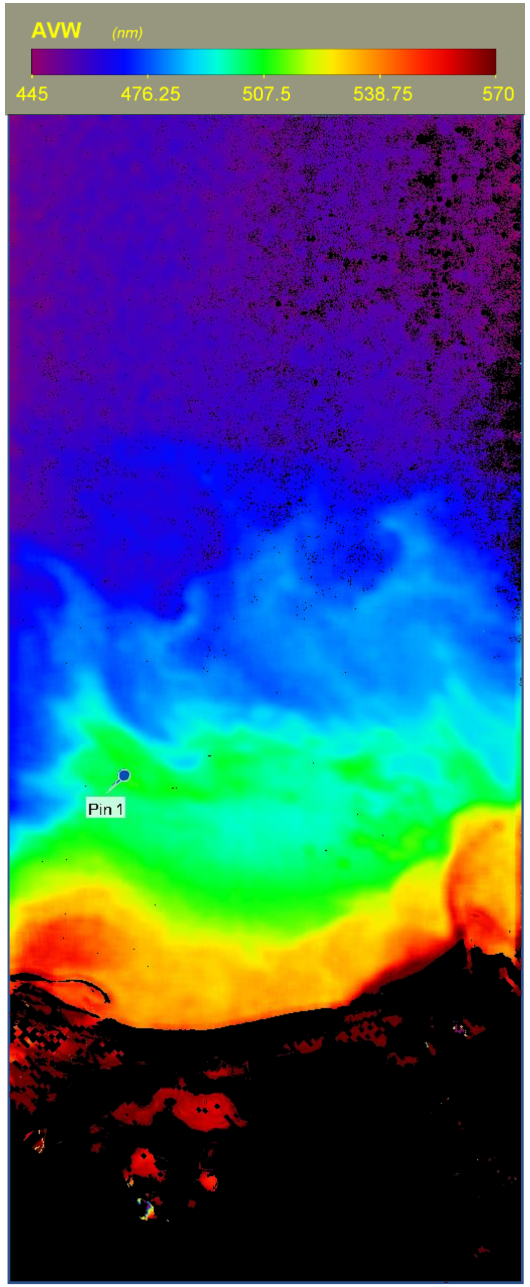



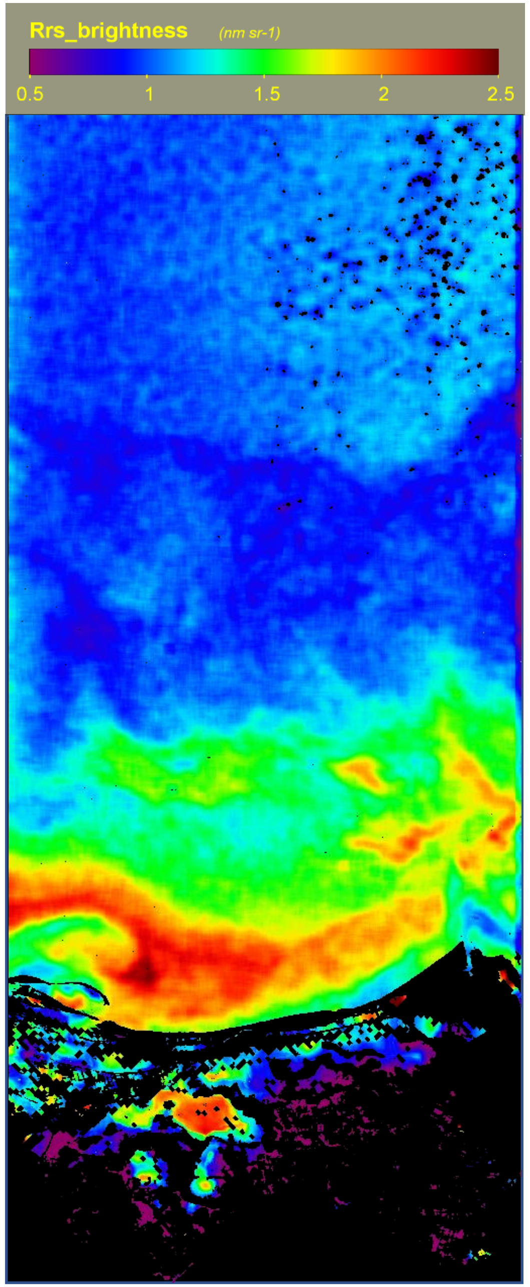
All $R_{rs}(\lambda)$ with same AVW



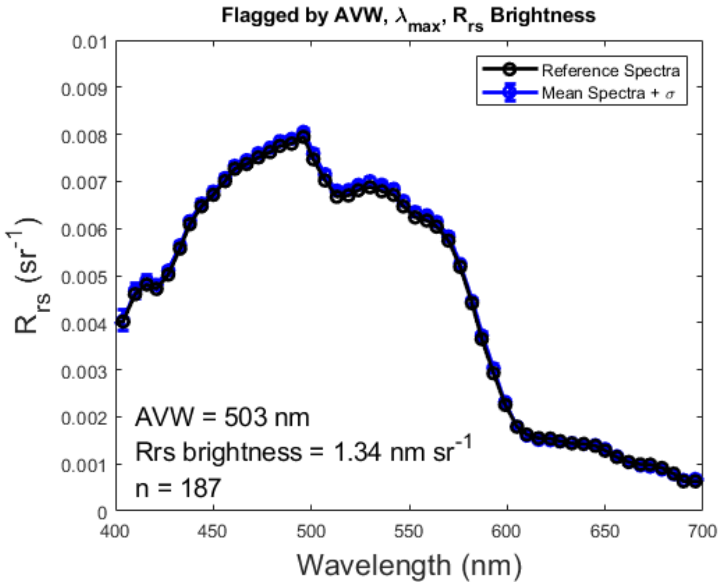

All $R_{rs}(\lambda)$ with same AVW, $R_{rs}$_brightness



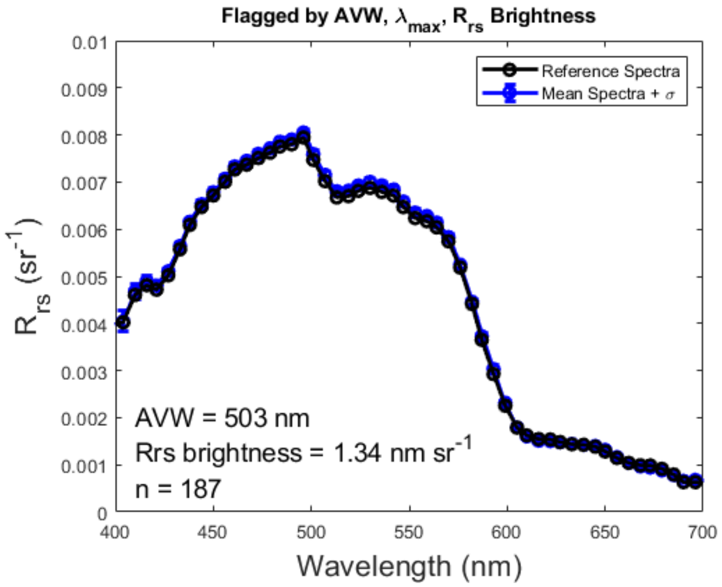



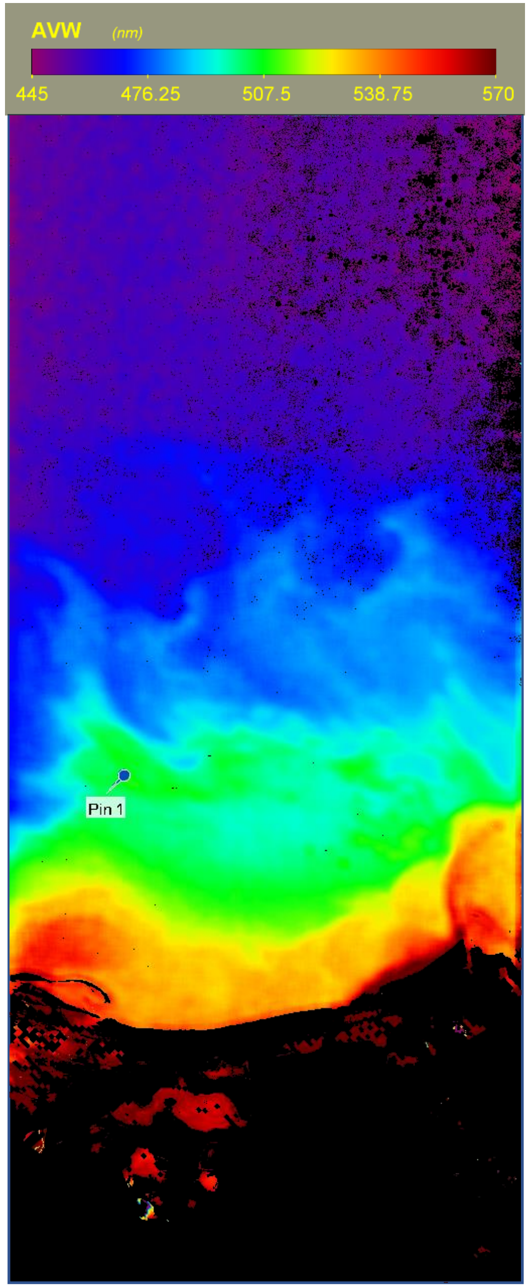



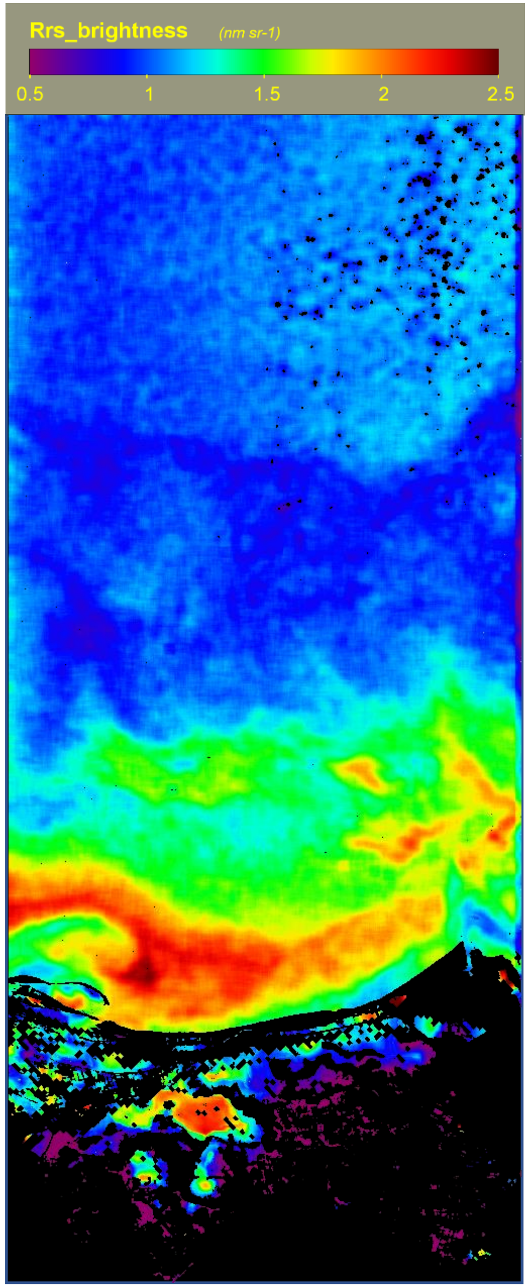
Adjust AVW ONLY

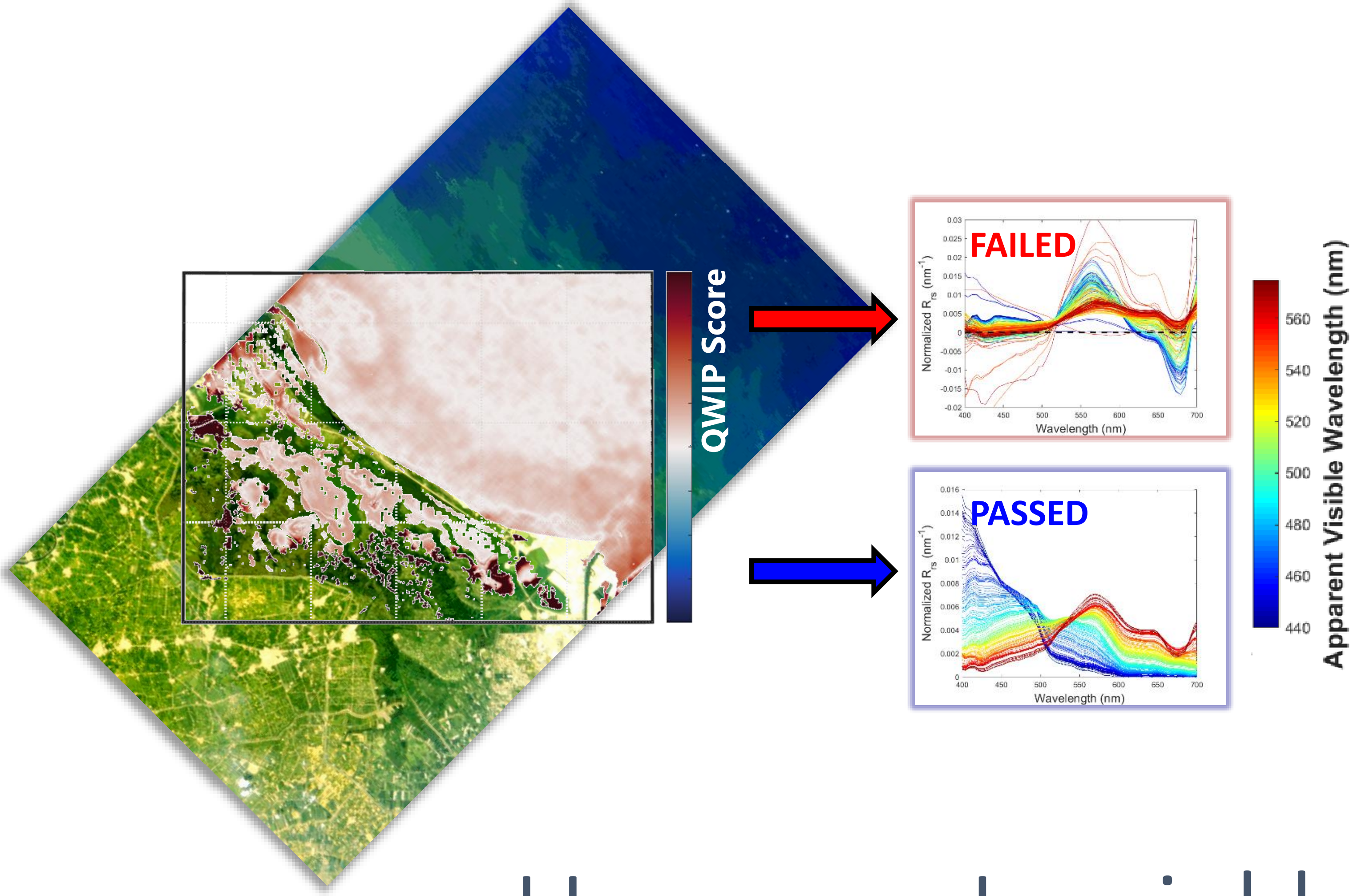



Adjust Rrs_Brightness ONLY



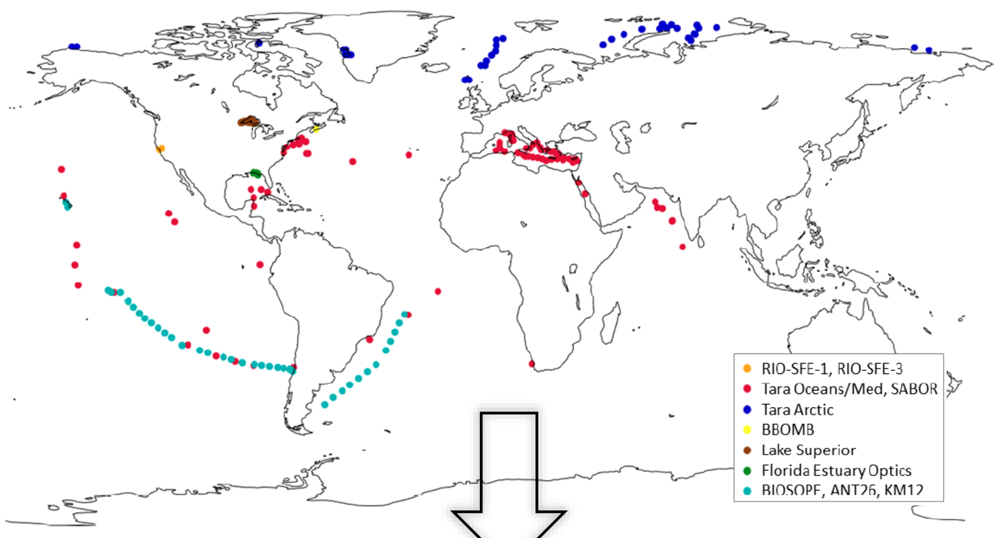



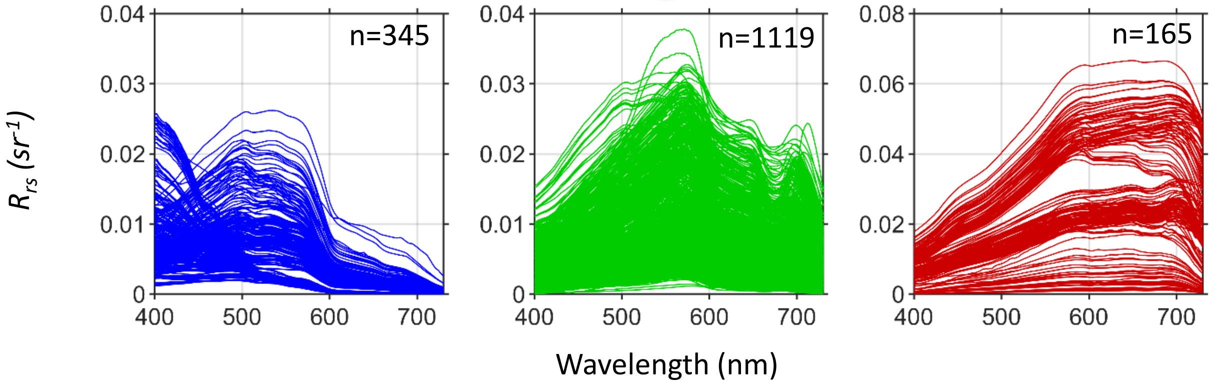



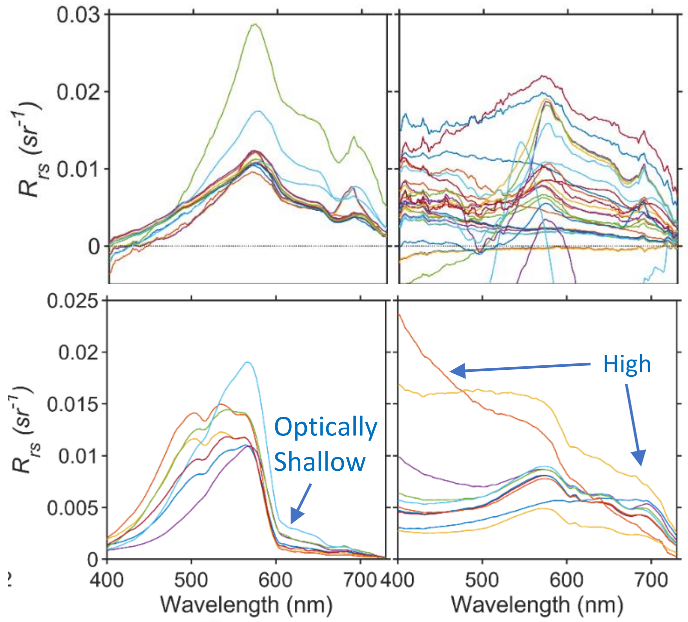



# How can I quickly assess the quality of my hyperspectral data?

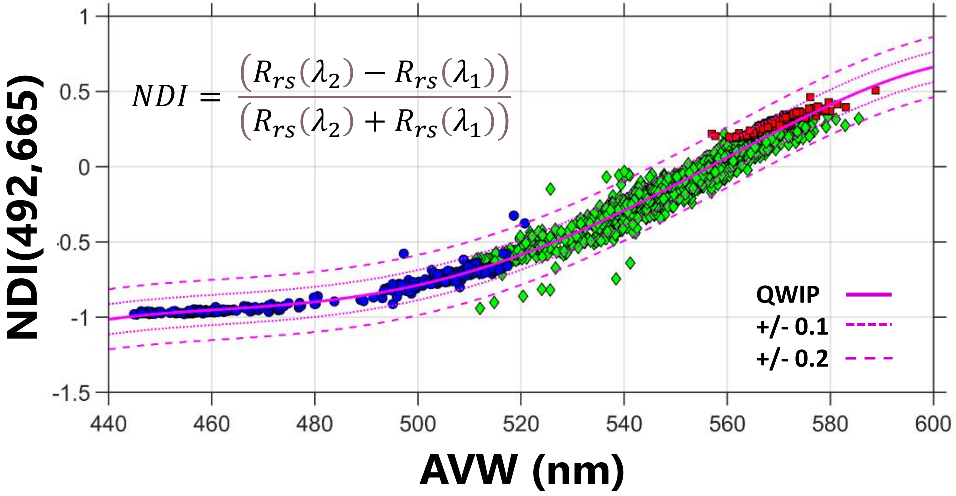
# Quality Water Index Polynomial (QWIP)



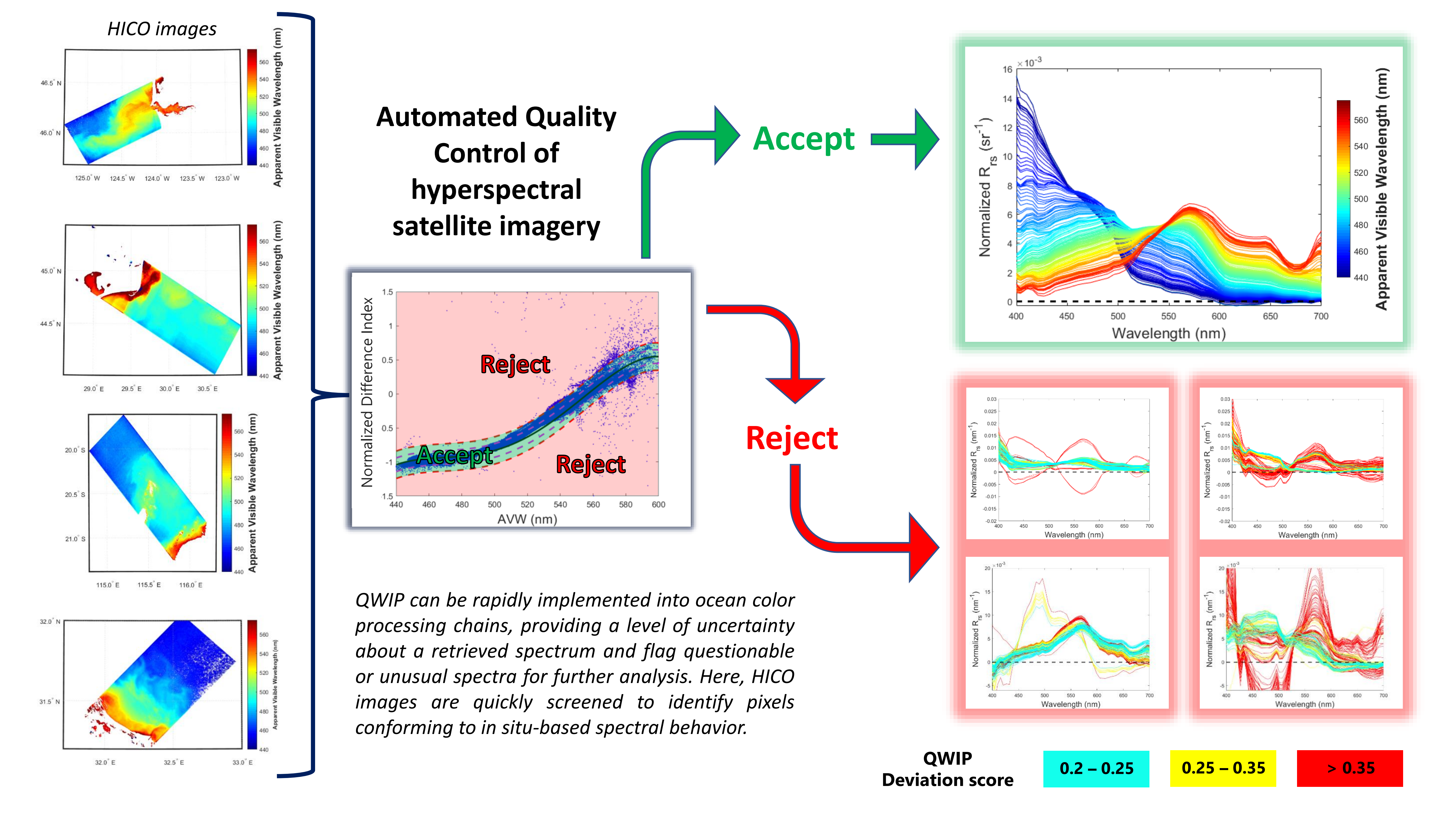




$R_{rs}$ data that deviate from QWIP are flagged

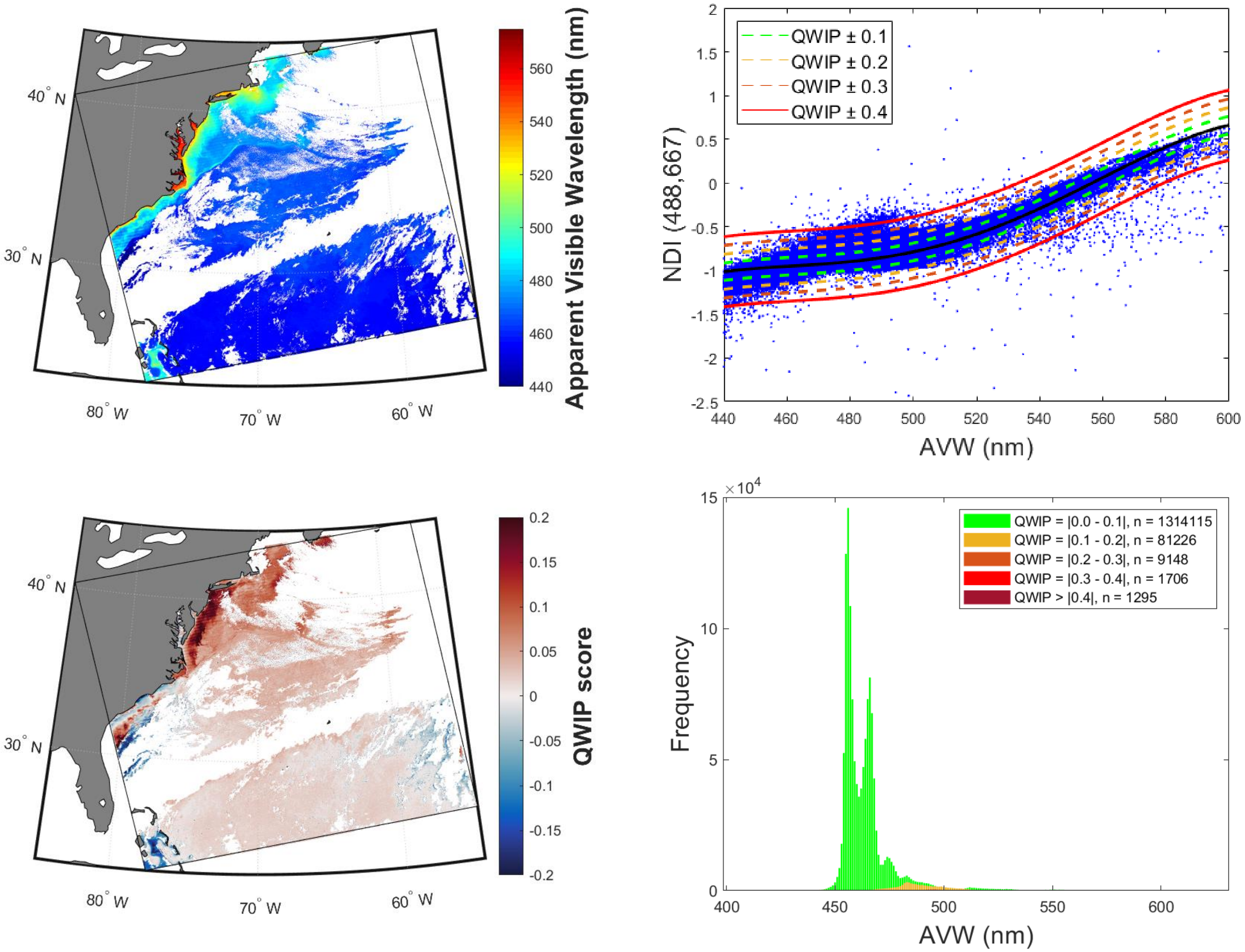



Quality control specifically developed for hyperspectral aquatic reflectance



HICO images
Apparent Visible Wavelength (nm)
Automated Quality Control of hyperspectral satellite imagery
Accept
Reject
Normalized Difference Index
AVW (nm)
Normalized $R_{rs}$ ($sr^{-1}$)
Wavelength (nm)
QWIP can be rapidly implemented into ocean color processing chains, providing a level of uncertainty about a retrieved spectrum and flag questionable or unusual spectra for further analysis. Here, HICO images are quickly screened to identify pixels conforming to in situ-based spectral behavior.
QWIP Deviation score
0.2 – 0.25
0.25 – 0.35
> 0.35

## MODIS Aqua - January 28, 2012

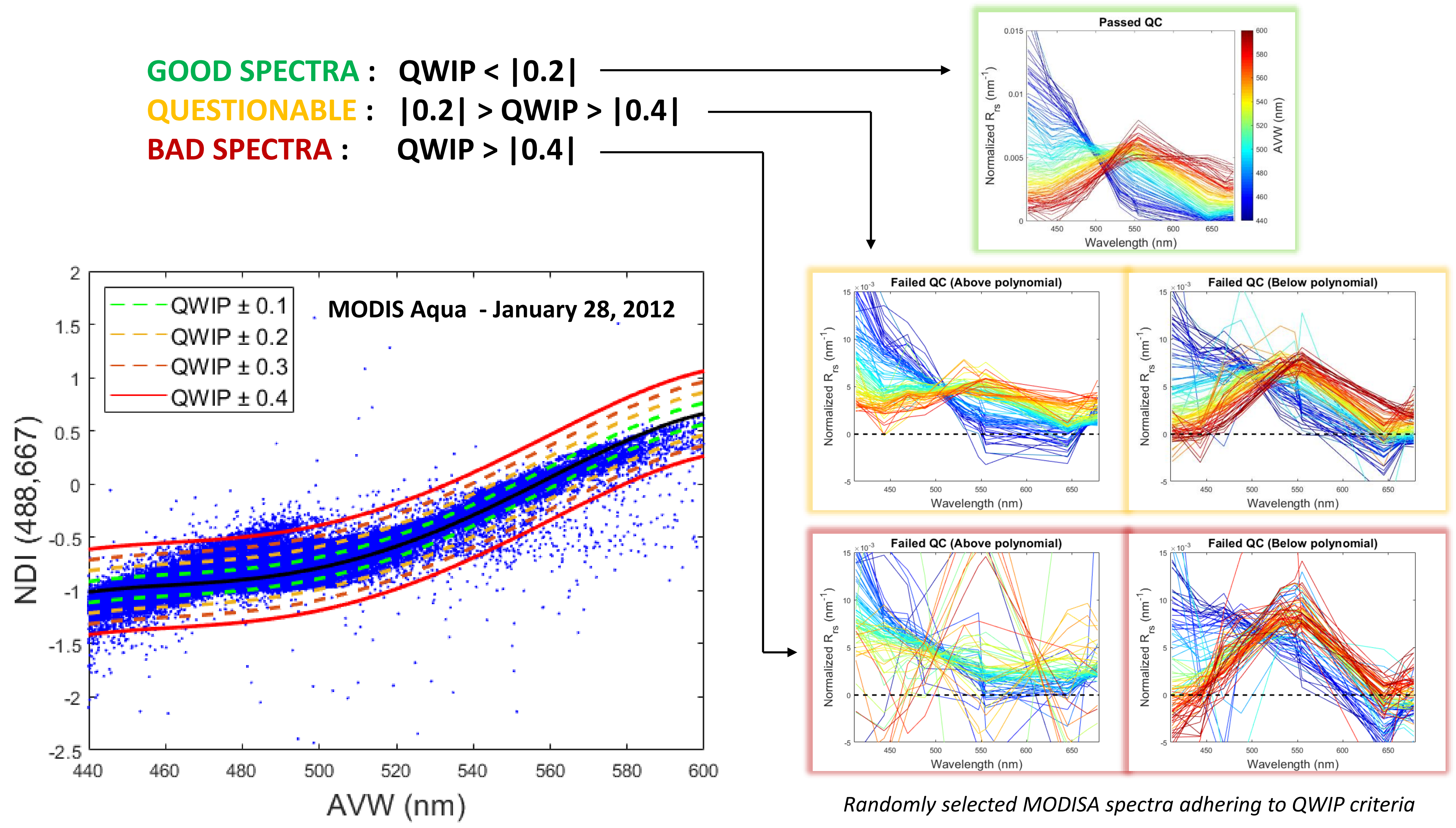



# Applying QWIP to multispectral data streams

Matlab code released with Dierssen et al. (2022) generates a series of informative plots for applying QWIP to hyperspectral and multispectral satellite scenes.

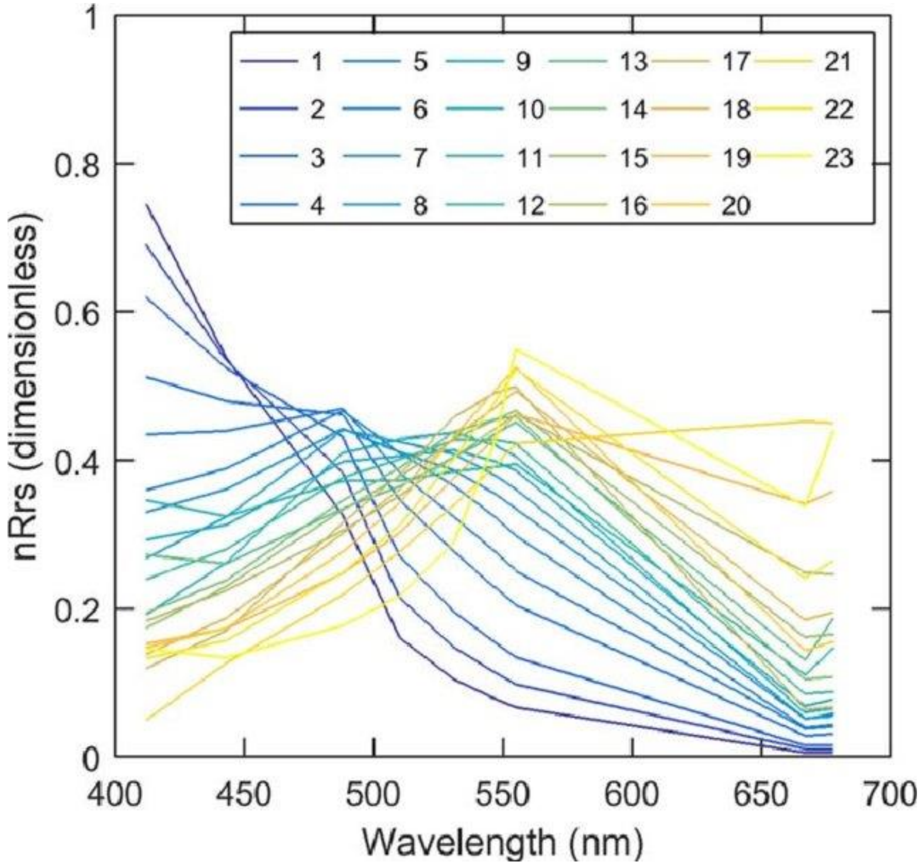
*Note*: Varying multispectral band placement generates biased AVW estimates, which must be corrected prior to applying QWIP. Bias correction detailed in ATBD: https://oceancolor.gsfc.nasa.gov/atbd/avw/.

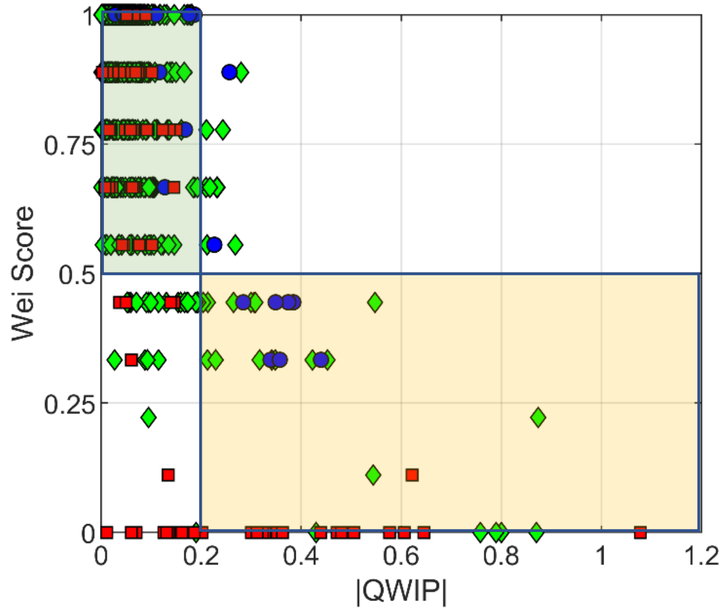
*Available for all sensors processed by NASA OBPG.*


GOOD SPECTRA : QWIP < |0.2|
QUESTIONABLE : |0.2| > QWIP > |0.4|
BAD SPECTRA : QWIP > |0.4|
MODIS Aqua - January 28, 2012
QWIP ± 0.1
QWIP ± 0.2
QWIP ± 0.3
QWIP ± 0.4
NDI (488,667)
AVW (nm)
Passed QC
Normalized R rs (nm-1)
Wavelength (nm)
AVW (nm)
Failed QC (Above polynomial)
Failed QC (Below polynomial)
Randomly selected MODISA spectra adhering to QWIP criteria

*QA score is determined as the number wavelengths for which the reflectance datum fit within the reference boundaries, divided by the total number of wavelengths assessed.*

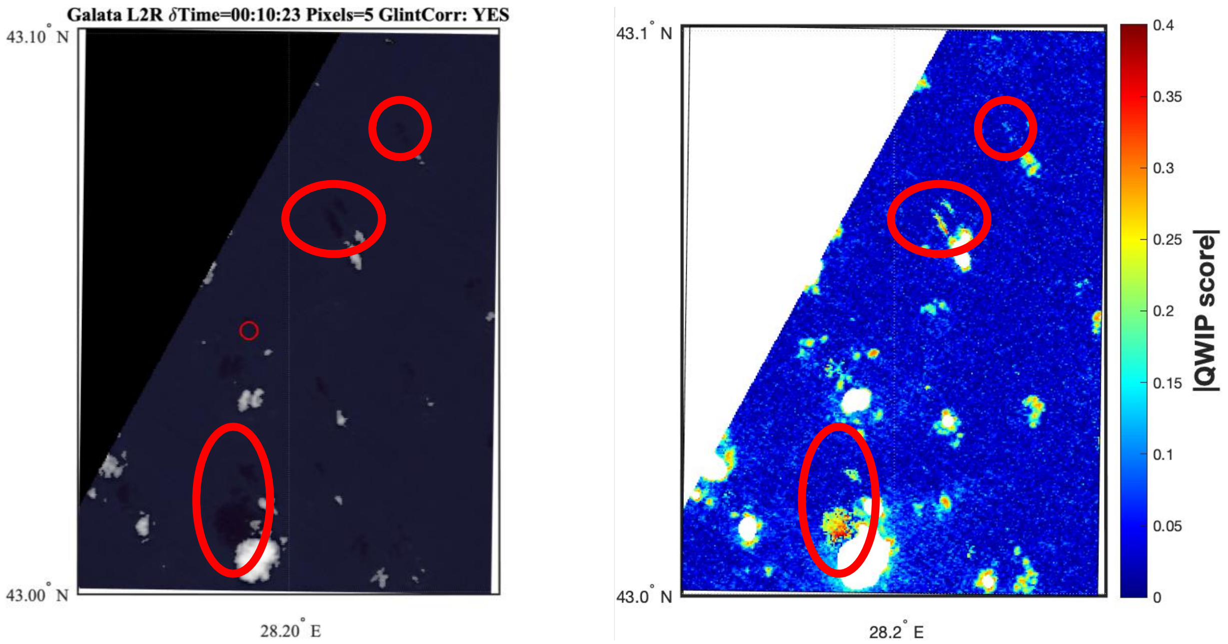
# QWIP v. QA Score (Wei et al. 2016)




| \|QWIP\| | Wei Score | |
|---|---|---|
| | Pass (>0.50) | Fail (<0.5) |
| **Pass (<0.2)** | 737<br>(B=25;G=641;Br=72)[a] | 39*<br>(B=0;G=23;Br=16) |
| **Fail (>0.2)** | 12<br>(B=2;G=10;Br=0) | 54<br>(B=8;G=23;Br=23) |

# QWIP detects cloud shadows in imagery



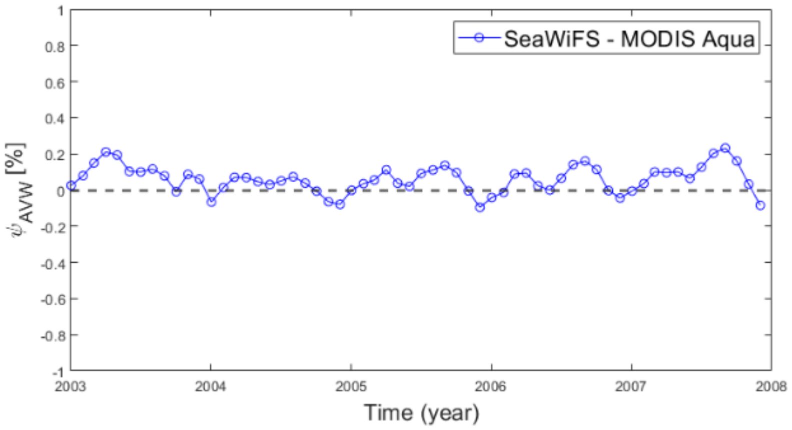

*Galata Platform (Black Sea) – DESIS HSI 2020-08-25*
*Images courtesy of Dirk Aurin*

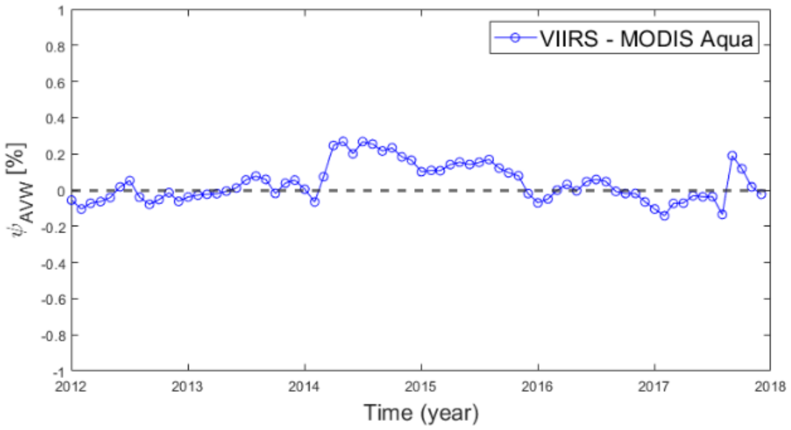
QWIP scores easily identify cloud shadows in high resolution DESIS imagery.

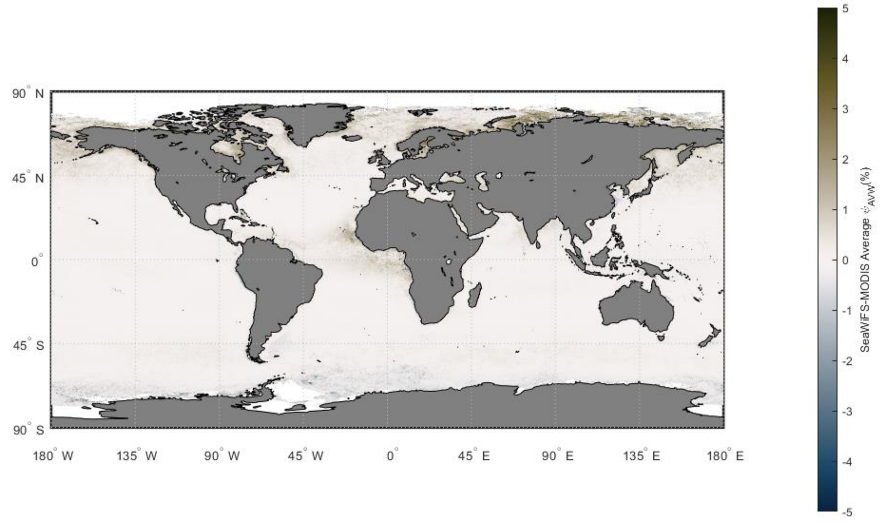
*QWIP may be potentially be adapted to TOA or Rayleigh-corrected spectra, providing a filter prior to fitting atmospheric correction aerosol models (TBD)…*

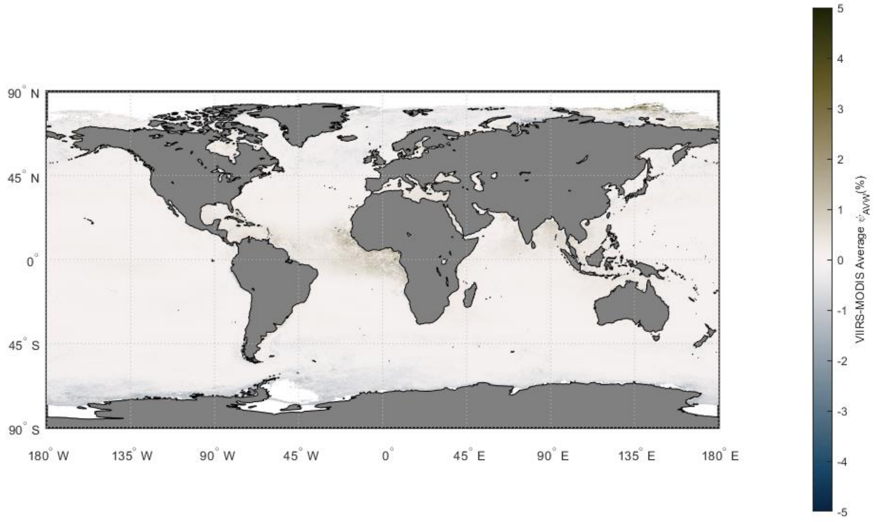
# Fitness-for-purpose, as featured in:

*Vandermeulen et al. (2020). 150 shades of green: Using the full spectrum of remote sensing reflectance to elucidate color shifts in the ocean. Remote Sensing of Environment, 247, 111900.*

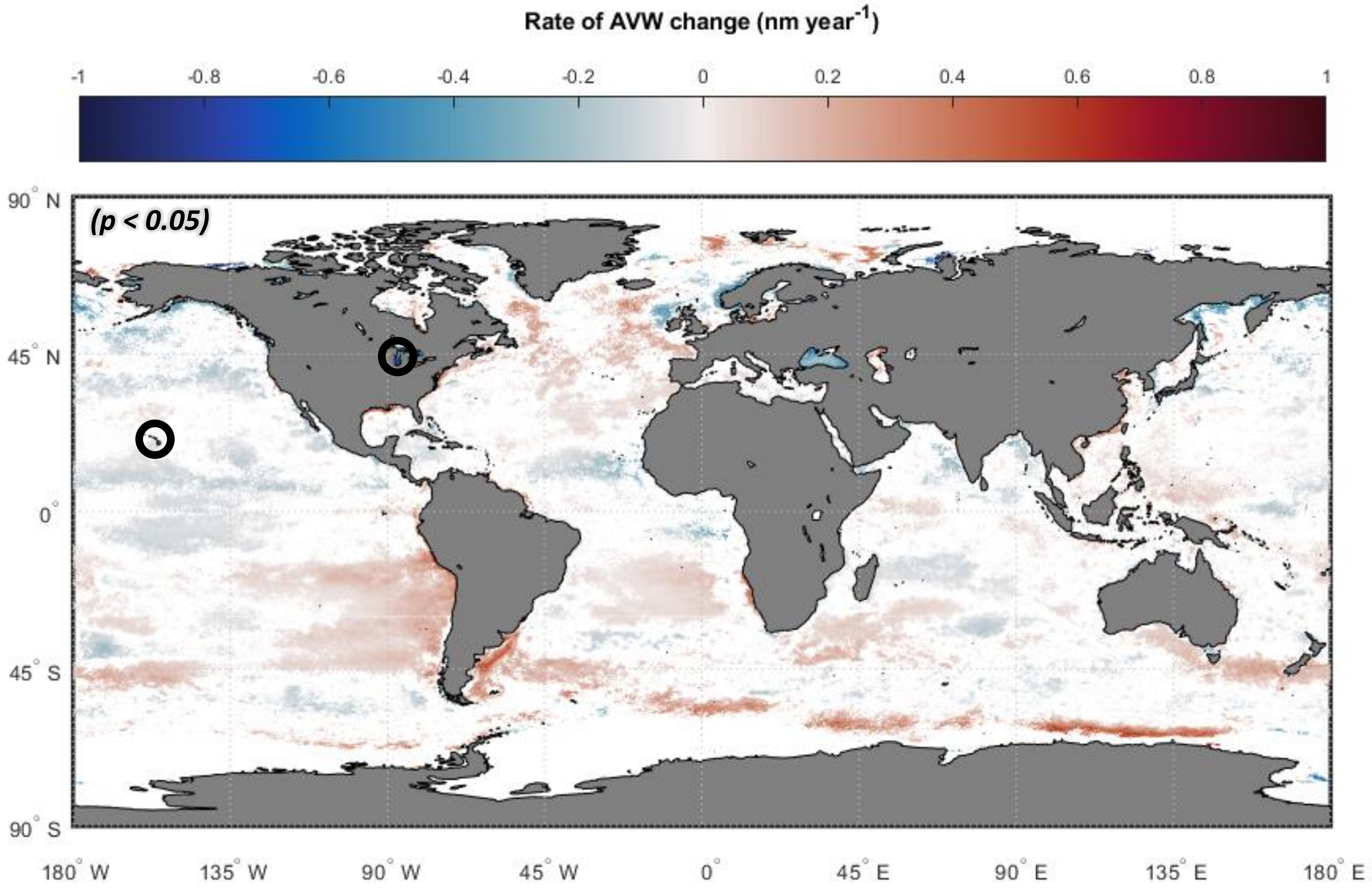


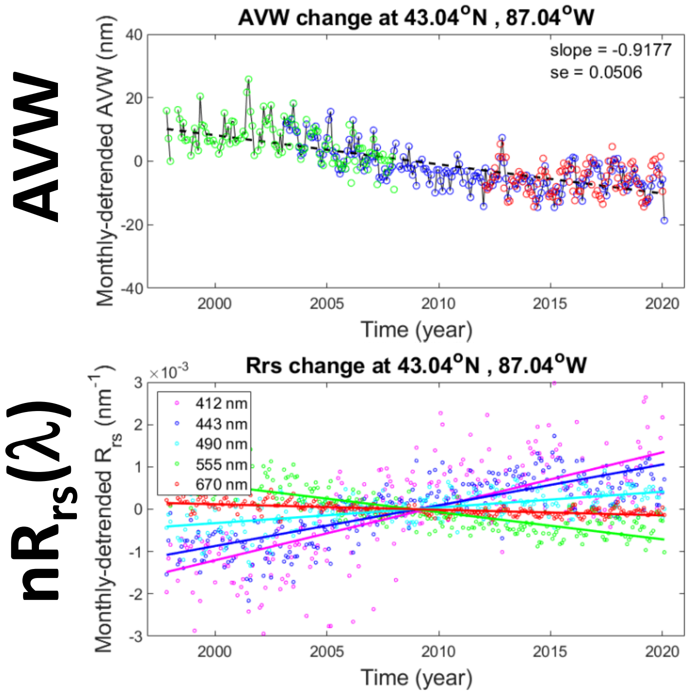



# Elucidating spectral shifts over time (SeaWiFS → MODIS/VIIRS → PACE)



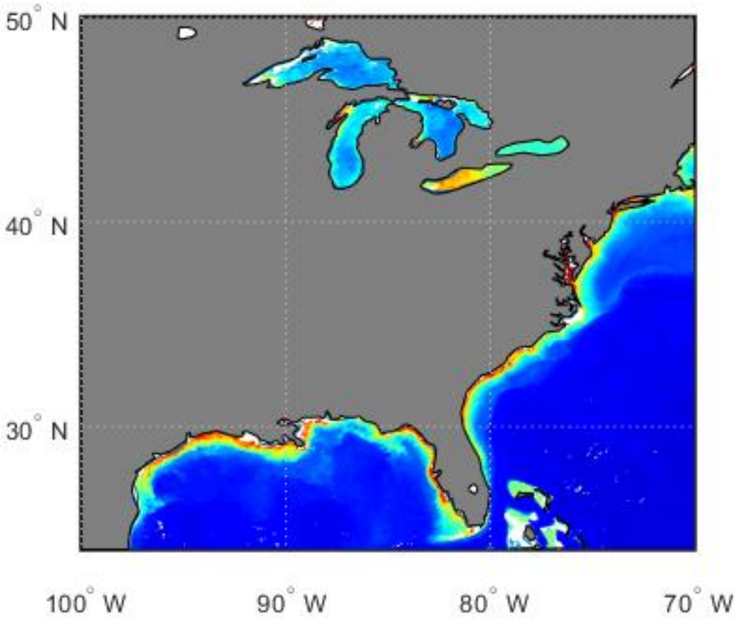



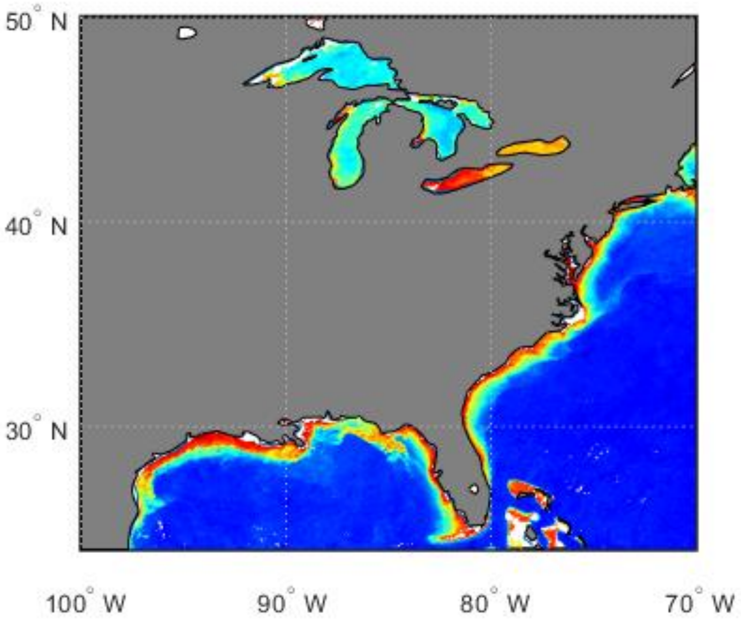

Robust Linear regression/time of AVW enables the examination of spectral shift in $R_{rs}(\lambda)$ over time

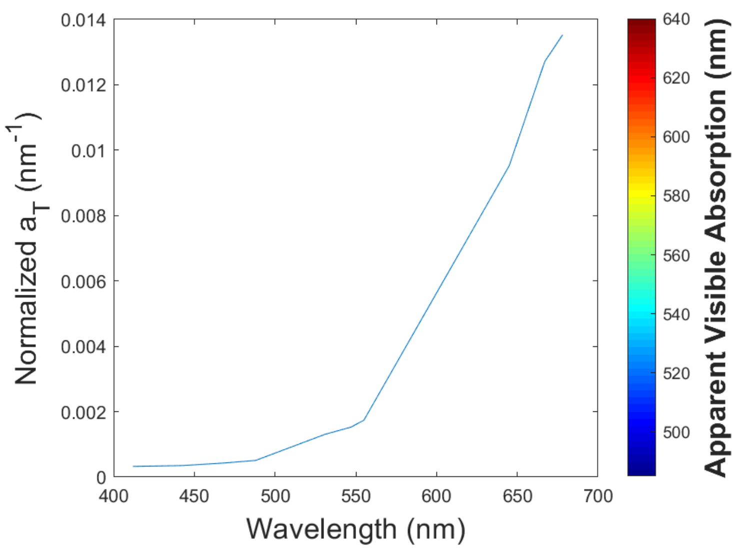
## Apparent Visible Wavelength

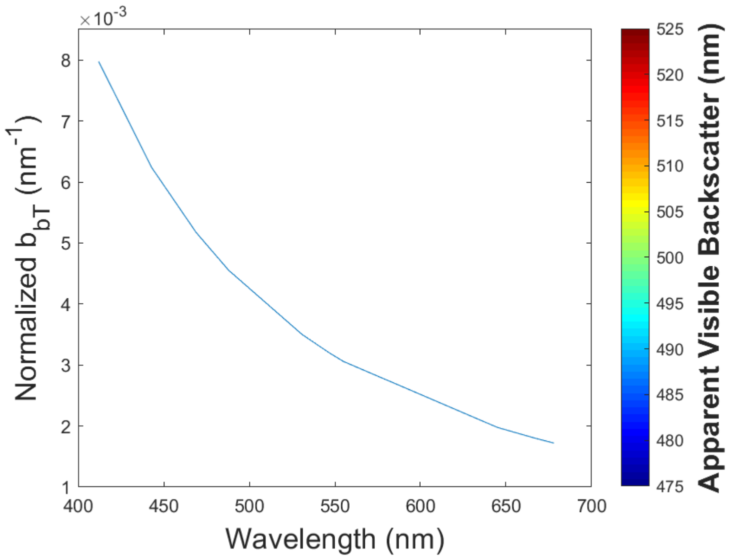
# ***into the IOP-verse***

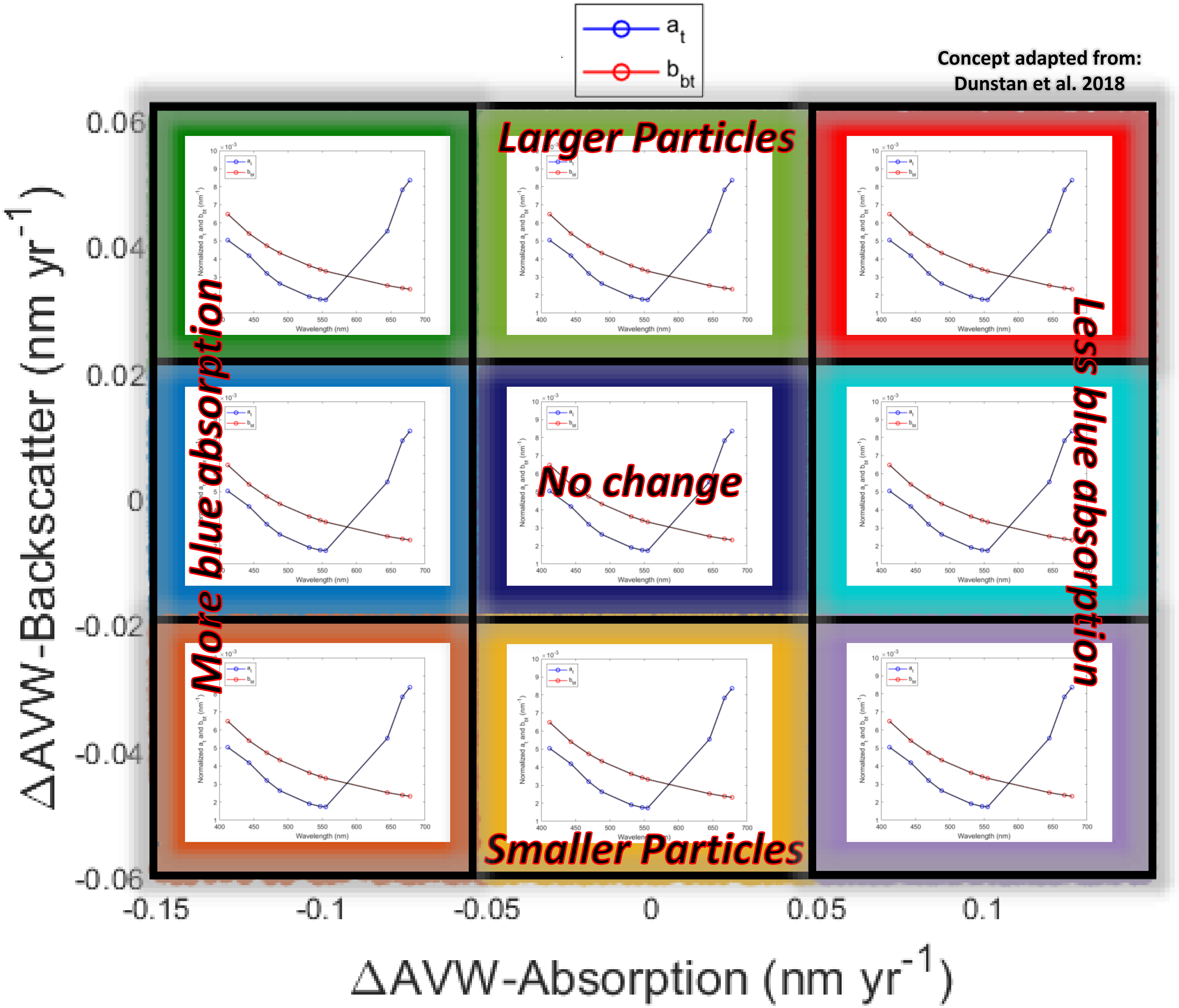
- AVW concept can be seamlessly applied to IOPs (total absorption and backscatter, GIOP).
- If we compare the rate of change in spectral shape of $a_T$ relative to that of $b_{bT}$, we find that the two are no always directly correlated or proportional.
- Subtle changes in the relative directionality of spectral trends may help elucidate bio-optical (or other) shifts within the satellite data record.



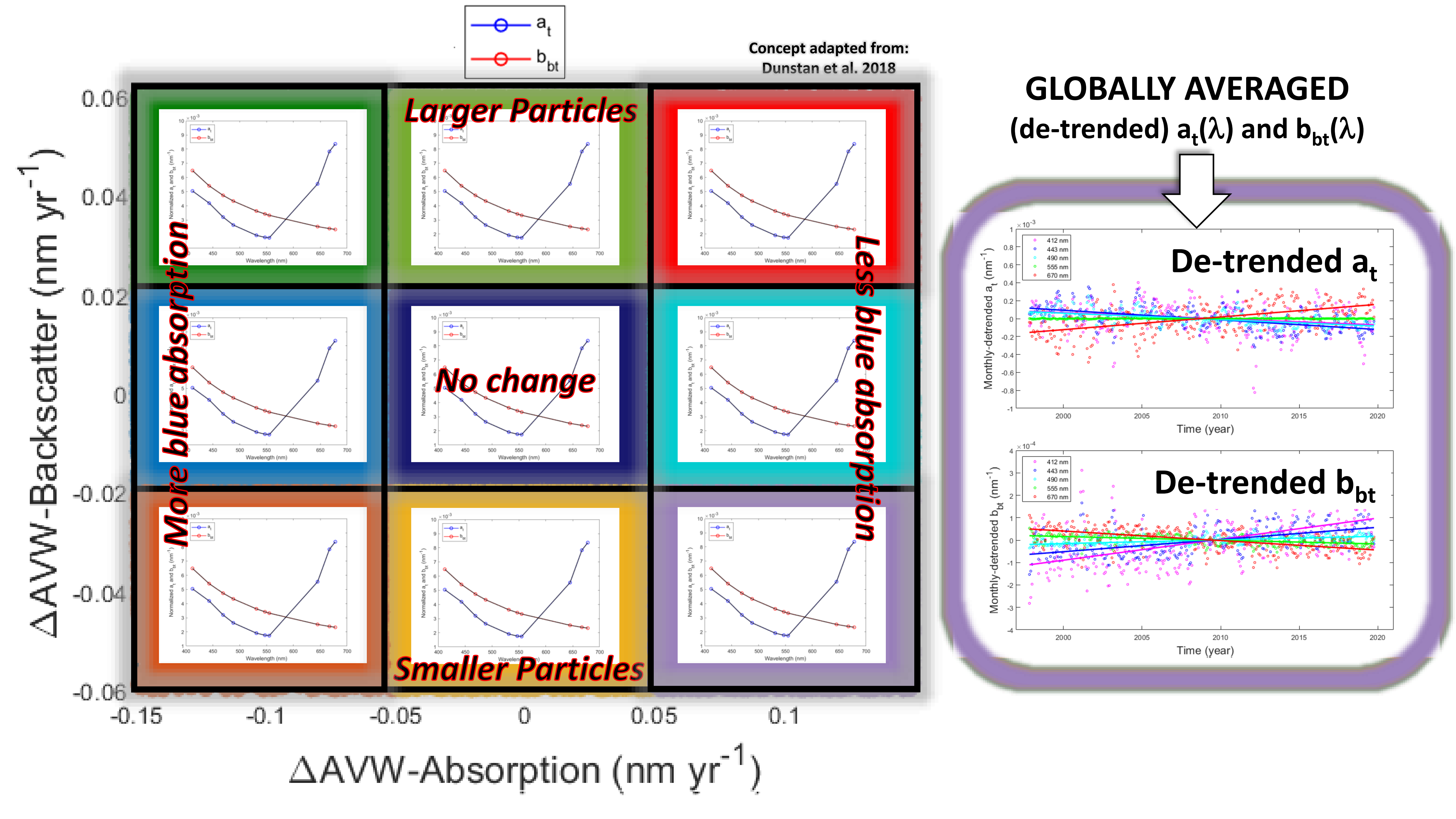






# Apparent Visible Wavelength *into the IOP-verse*

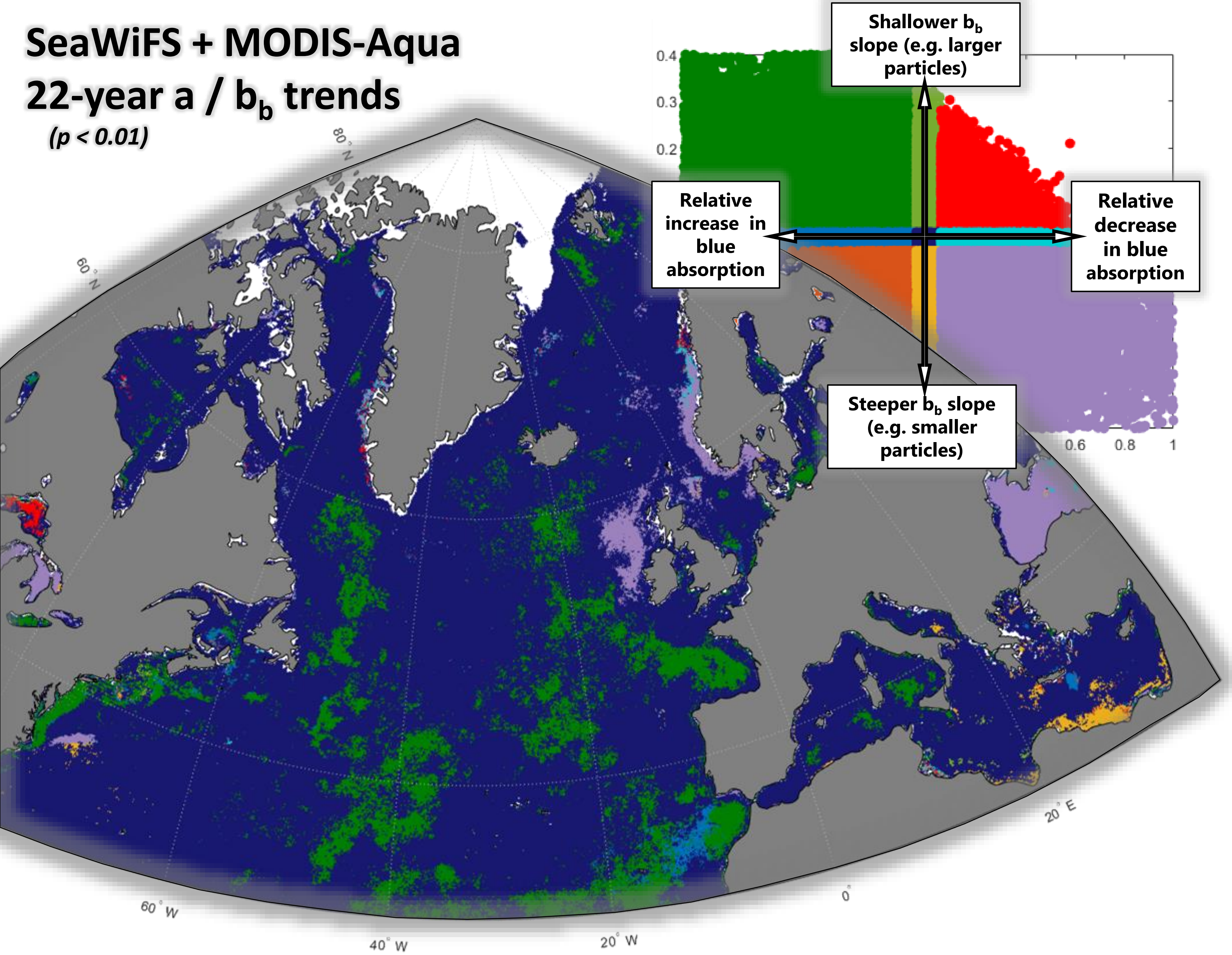
- AVW concept can be seamlessly applied to IOPs (total absorption and backscatter, GIOP).
- If we compare the rate of change in spectral shape of $a_T$ relative to that of $b_{bT}$, we find that the two are no always directly correlated or proportional.
- Subtle changes in the relative directionality of spectral trends may help elucidate bio-optical (or other) shifts within the satellite data record.



a_t
b_bt
Concept adapted from:
Dunstan et al. 2018
GLOBALLY AVERAGED
(de-trended) a_t(λ) and b_bt(λ)
Larger Particles
No change
Smaller Particles
More blue absorption
Less blue absorption
ΔAVW-Backscatter (nm yr^-1)
ΔAVW-Absorption (nm yr^-1)
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.15
-0.1
-0.05
0
0.05
0.1
Wavelength (nm)
Normalized a_t and b_bt (nm^-1)
De-trended a_t
De-trended b_bt
412 nm
443 nm
490 nm
555 nm
670 nm
Monthly-detrended a_t (nm^-1)
Monthly-detrended b_bt (nm^-1)
Time (year)
2000
2005
2010
2015
2020

# SeaWiFS + MODIS-Aqua 22-year a / $b_b$ trends

*($p < 0.01$)*




## What this does show:

- Relative *magnitude* and *directionality* of spectral shifts in GIOP products.

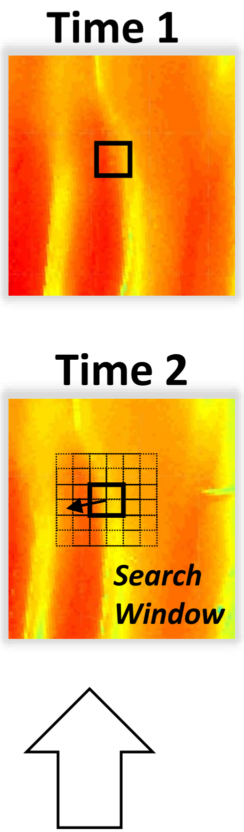
## What this does NOT show:

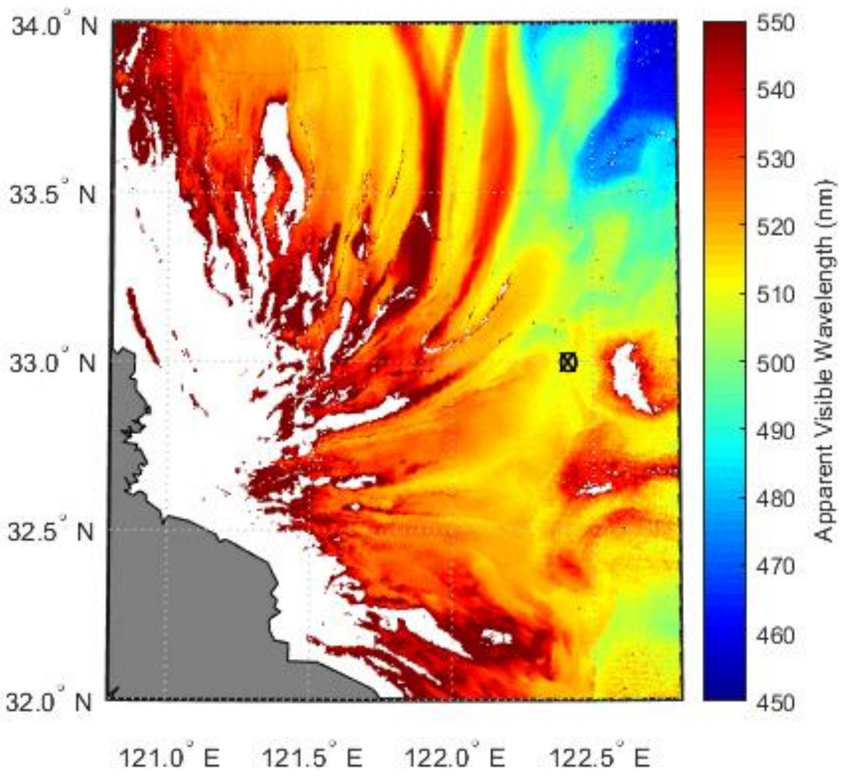
- Where these measurements have been adversely impacted by atm-correction.
- A truly independent assessment of IOP behavior ($S = 0.018$, $\eta \sim r_{rs}443/r_{rs}55x$, $a_{ph} \sim$ chl-a $\sim r_{rs}443/r_{rs}55x$).

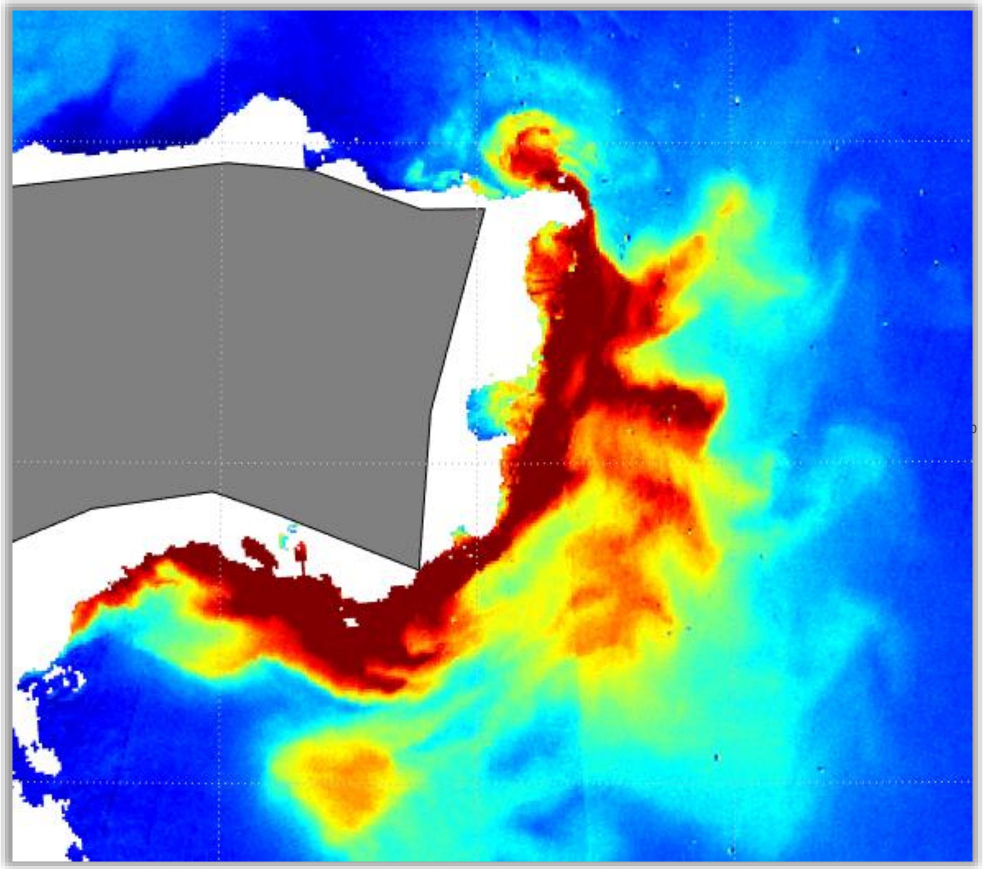
## So why do we care?

- The simultaneous conceptualization of multiple data dimensions maximizes the full capacity of a given sensor. Can we link bio-optical behavior over time with shifts in community composition?

# Surface fluxes derived from integrated Ocean Color products



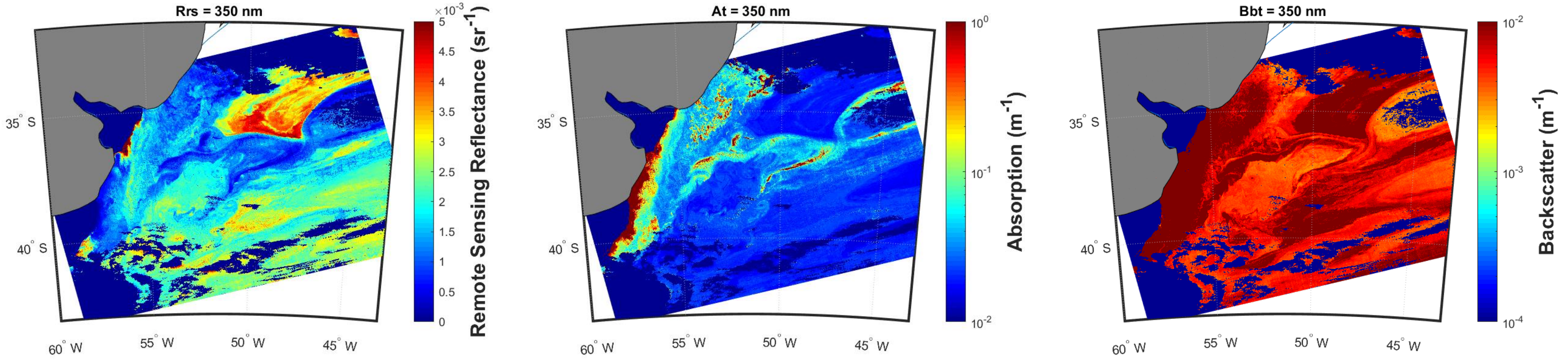



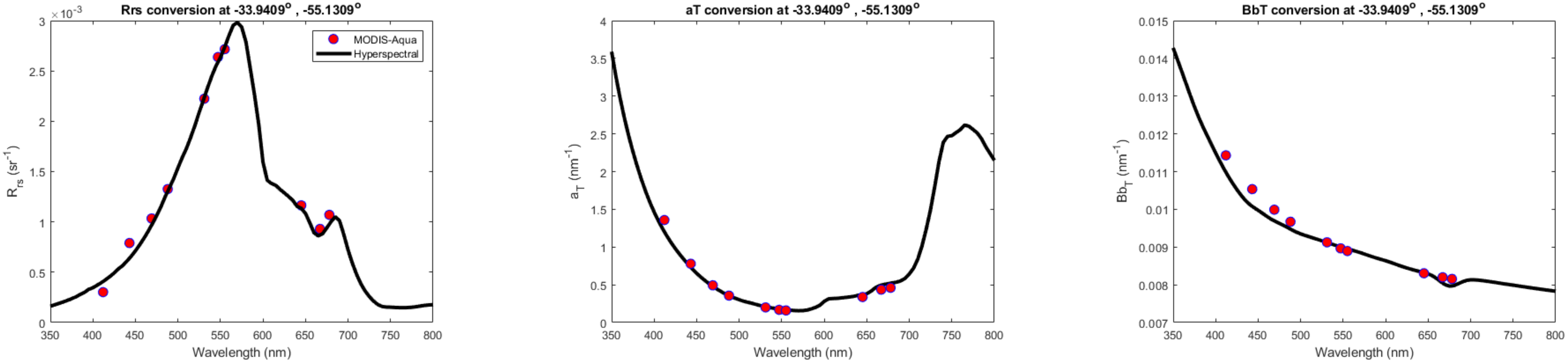



$$MCC = \frac{\sum_i \sum_j [\varphi_1 \varphi_2]}{\sqrt{\sum_i \sum_j [\varphi_1]^2 \sum_i \sum_j [\varphi_2]^2}}$$

*Changing ocean color products from consecutive images can be used to estimate surface currents using the **Maximum Cross Correlation (MCC)** method. MCC is based on determining the similarity in the spatial gradients of ocean color between multiple images. The surface currents are determined based on the movement of a spatial gradient of the Apparent Visible Wavelength product from one hour to next.*

# Through empirical tuning of multispectral radiometry and IOPs, synthetic data can be effortlessly mapped to MODIS coordinates

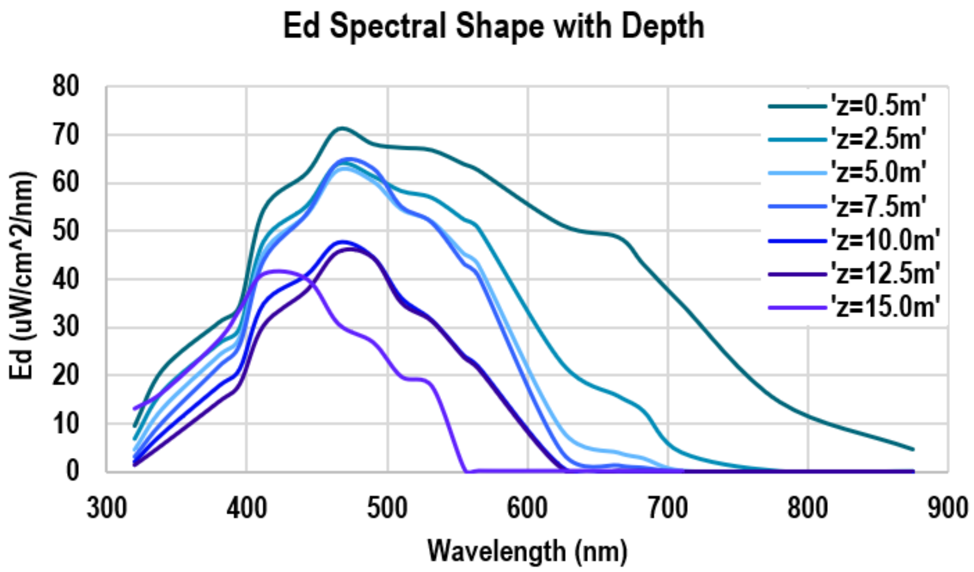


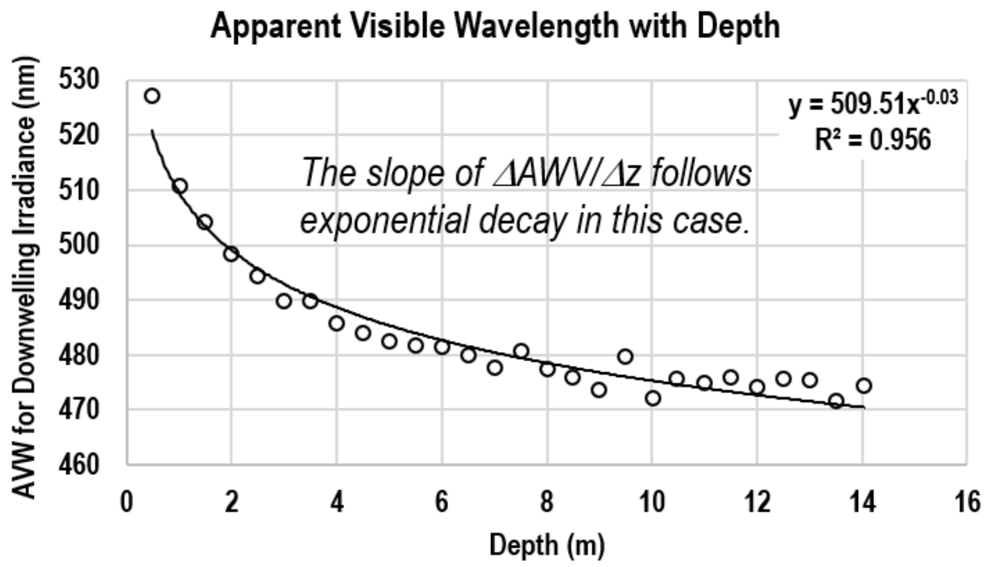



# $\Delta$AVW/$\Delta$z

- Changes in spectral shape with depth can be defined by an exponential decay function, in a manner similar to $K_d$.
- The slope of $\Delta$AVW/$\Delta$z ($K_{AVW}$) is an additional classification which directly describes the quality of light available in the water column.
- Spectral $K_d$ information from GIOP/QAA may enable global reconstruction of $K_{AVW}$.


Lee, Z., Shang, S., Li, Y., Luis, K., Dai, M., & Wang, Y. (2021). Three-Dimensional Variation in Light Quality in the Upper Water Column Revealed With a Single Parameter. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-10.

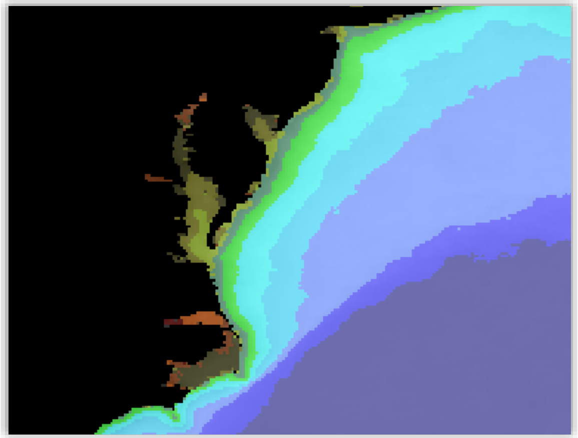



# Parting thought…

*In a recent study, Dutkiewicz et al. (2019) suggested that, $R_{rs}(\lambda)$, defined as the light that exits the water column and is measured by satellite-borne instruments, will provide the earliest and strongest signal of marine ecosystem change because they are not subject to the natural variability and algorithm uncertainties as derived products. Moreover, responses in $R_{rs}(\lambda)$ not only include the signal of Chlorophyll-a, but all of the optically active constituents in the water, which include colored dissolved organic matter (CDOM), nonalgal particles, and phytoplankton pigment signatures, all of which influence the light field of the water column.*

# Applications:



- Holistic identification of optical water types/seascapes
- Analysis of spectral variance for targeted sampling
- Potential improvements to semi-analytical inversions
- A useful climatological metric of change
- Correlation of similar water types on global scales
- Useful for display and analysis of multi/hyperspectral *in situ* data
- Identification of phytoplankton community characteristics
- Implementation of decision tree approaches for algorithm development
- Quality control check of algorithm performance (e.g. erratic spectral shapes)
- Spectrum matching and search functions

# Summary – *Use full spectral information*

1) The weighted harmonic mean of spectral wavelengths can be used as an index of optical water types that share similar spectral signatures.
2) This technique can be translated to relate spectral shape between sensors of varying spectral resolution.
3) The output along a continuum of color values enables spatio-temporal analysis of subtle spectral shifts.
4) Utilizing paired full spectral metrics enables the efficient and effective assessment of data quality.
5) Conceptualization of multiple data dimensions is imperative as passive/active sensors grow increasingly sophisticated in nature.