On the data quality and quantity of VIIRS/SNPP ocean color data products: from research to applications

Chuanmin Hu, University of South Florida
huc@usf.edu, https://optics.marine.usf.edu

Co-authors and collaborators:
Brian Barnes, Mengqiu Wang, Jennifer Cannizzaro, David English, Shaojie Sun, Lin Qi, Lian Feng, Paul Carlson, Tracey Sutton, Menghua Wang, Lide Jiang, Michael Ondrusek……

Outline

1. Data quality: standard and non-standard data products
2. Data quantity: not always a zero-sum game
3. Examples on research and applications: algal blooms (harmful and non-harmful, microalgae and macroalgae) and oil spills

VIIRS data sources:

- NOAA/NESDIS/STAR (MSL12 processing)
- NASA GSFC (L2GEN processing)

MODIS data source:

- NASA GSFC (L2GEN processing)

1. Data quality: validation using field measurements

Table 1: Summary of selected *Rrs* validation methods and results

Citation	Platform	Environment	Sensor	Processing*	CV Threshhold (box size)	Temporal overlap (hr)	Accuracy Statistic (547 or 551 nm)
Mélin et al., 2007	Fixed	Coastal	MODIS	SeaDAS 5.0 (~2005.1)	0.2 (3x3)	3.5	MAPD = 14%
Antoine et al., 2008	Fixed	Oceanic	MODIS	SeaDAS 2005.0	- (5x5)	3	MAPD = 17%
Zibordi et al., 2009	Fixed (AERONET)	Coastal	MODIS	SeaDAS 5.2 (~2005.1)	0.2 (3x3)	2	MAPD = 10%
Maritorena et al., 2010	Ship & Fixed (AERONET)	Coastal & Oceanic	MODIS	SeaDAS 2005.1	- (-)	-	MR = 1.006
Hlaing et al., 2013	Fixed (AERONET)	Coastal	MODIS	SeaDAS 2012.0	0.2 (3x3)	2	MAPD ~ 12%
Brando et al., 2016	Ship & Fixed (AERONET)	Coastal & Oceanic	MODIS	SeaDAS 2014.0.1	- (3x3)	2	MAPD ~ 12%
Wang et al., 2013	Fixed (MOBY)	Oceanic	VIIRS	MSL12	- (11x11)	-	MR = 0.98
Hlaing et al., 2013	Fixed (AERONET)	Coastal	VIIRS	SeaDAS 2012.2	0.2 (3x3)	2	MAPD ~ 12%
Hlaing et al., 2013	Fixed (AERONET)	Coastal	VIIRS	MSL12 (IDPS v6.6)	0.2 (3x3)	2	MAPD = 14%
Ahmed et al., 2013	Fixed (AERONET)	Coastal	VIIRS	SeaDAS 2013.0	0.2 (3x3)	2	MAPD = 10 - 15%
Brando et al., 2016	Fixed (AERONET)	Coastal	VIIRS	MSL12	0.2 (3x3)	2	MAPD = 14%
Wang et al., 2014	Fixed (MOBY)	Oceanic	VIIRS	MSL12	- (5x5)	8	MR = 0.992
Vandermeulen et al., 2015	Ship & Fixed (AERONET)	Coastal & Oceanic	VIIRS	NRL-APS v5.1	- (-)	3	RMSE = 0.160 mW/cm^2/m/sr
Wang et al., 2015	Fixed (MOBY)	Oceanic	VIIRS	MSL12	- (5x5)	-	MR = 1.0157
Wang et al., 2016	Fixed (MOBY)	Oceanic	VIIRS	MSL12	- (5x5)	-	MR = 1.0148
Brando et al., 2016	Ship & Fixed (Aeronet)	Coastal & Oceanic	VIIRS	SeaDAS 2014.0.1	- (3x3)	2	MAPD ~ 12%

MR = Mean Ratio, MAPD = Mean Absolute Percent Difference, RMSE = Root Mean Squared Error, - = not performed or not reported, * Where not specified, approximate processing version reported

Table 2: Level-2 Processing Flags (from http://oceancolor.gsfc.nasa.gov/atbd/ocl2flags/ and Wang et al., 2017).

Bit position	Default Mask	L3 Mask*	"Current" Mask	Bailey and Werdell (2006) †	Hlaing et al. (2013) §	Name (NASA)	Name (NOAA)	NASA Description [NOAA description]
0		X	X	X	X	ATMFAIL	ATMFAIL	Atmospheric correction failure
1	X	X	X	X	X	LAND	LAND	Pixel is over land
2						PRODWARN	PRODWARN	Warning from ≥ 1 product algorithms
3		X	X	X	X	HIGLINT	HIGLINT	Sunglint: reflectance exceeds threshold
4	X	X	X	X	X	HILT	HILT	Radiance very high or saturated
5		X	X	X	X	HISATZEN	HISATZEN	Sensor zenith angle exceeds threshold
6						COASTZ	COASTZ	Pixel is in shallow water
7						Spare	LANDADJ	[Probable land-adjacent contamination]
8		X	X	X	X	STRAYLIGHT	STRAYLIGHT	Probable stray light contamination
9	X	X	X	X	X	CLDICE	CLOUD	Probable cloud or ice contamination
10		X				COCCOLITH	COCCOLITH	Coccolithophores detected
11						TURBIDW	TURBIDW	Turbid water
12		X	X	X	X	HISOLZEN	HISOLZEN	Solar zenith angle exceeds threshold
13						Spare	HITAU	[High Aerosol Optical Thickness]
14		X	X	X	? §	LOWLW	LOWLW	Very low water-leaving radiance
15		X		? †		CHLFAIL	CHLFAIL	Chlorophyll algorithm failure
16		X	X		?	NAVWARN	NAVWARN	Navigation quality is suspect
17		X				ABSAER	ABSAER	Absorbing Aerosols determined
18						Spare	CLDSHDSTL	[Cloud straylight or shadow]
19		X	X			MAXAERITER	MAXAERITER	NIR iteration limit reached
20				?	X	MODGLINT	MODGLINT	Moderate sun glint
21				? †		CHLWARN	CHLWARN	Chlorophyll out-of-bounds
22		X	X			ATMWARN	ATMWARN	Atmospheric correction is suspect
23						Spare	ALGICE	[Sea ice identified by nLw]
24						SEAICE	SEAICE	Pixel is over sea ice
25		X	X		X	NAVFAIL	NAVFAIL	Navigation failure
26						FILTER	FILTER	Insufficient data for smoothing filter
27						Spare	ALTCLD	[Cloud detected]
28						BOWTIEDEL	FOG	VIIRS deleted overlapping pixels [Fog]
29						HIPOL	FROMSWIR	High polarization [SWIR atm. corr. used]
30						PRODFAIL	PRODFAIL	Failure in any product
31						SPARE	OCEAN	[Pixel is over ocean]

* The L3 mask is used for generation of global composite data products.

† Includes additional flag(s) specific to C_a. Also used by Antoine et al. (2008), Mélin et al. (2007), Zibordi et al. (2009)

§ Includes additional flag for negative Rayleigh-corrected reflectance. Also used by Ahmed et al. (2013).

Validation using data collected in North America

(most data collection supported by NOAA VIIRS cal/val program)

Station locations and partitions

VIIRS NOAA: MSL12 Apr 2017 SDR
VIIRS NASA and MODISA: L2GEN r2018.0

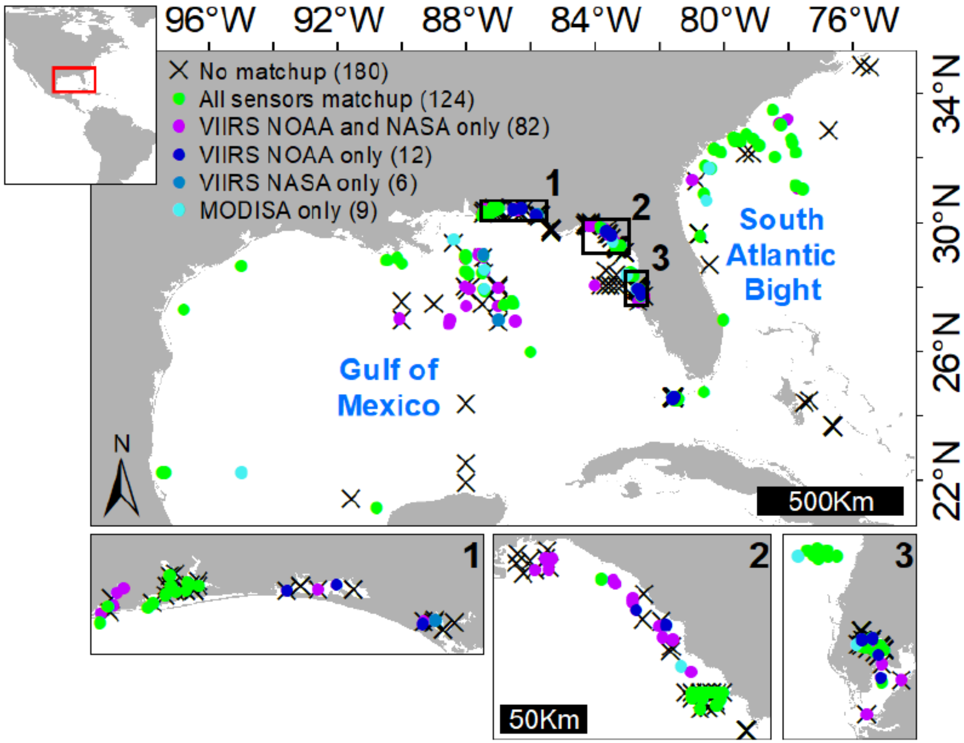
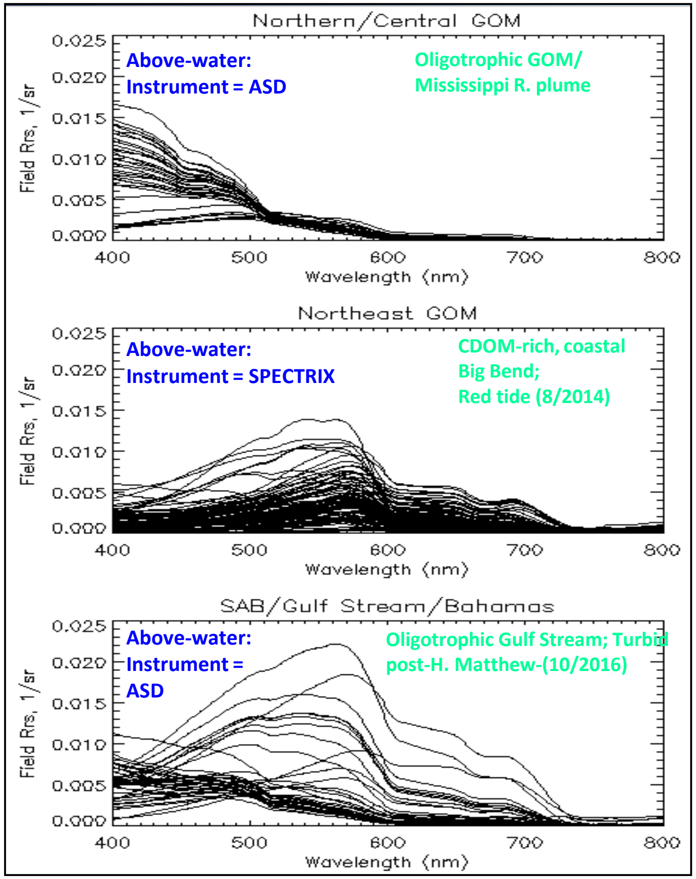


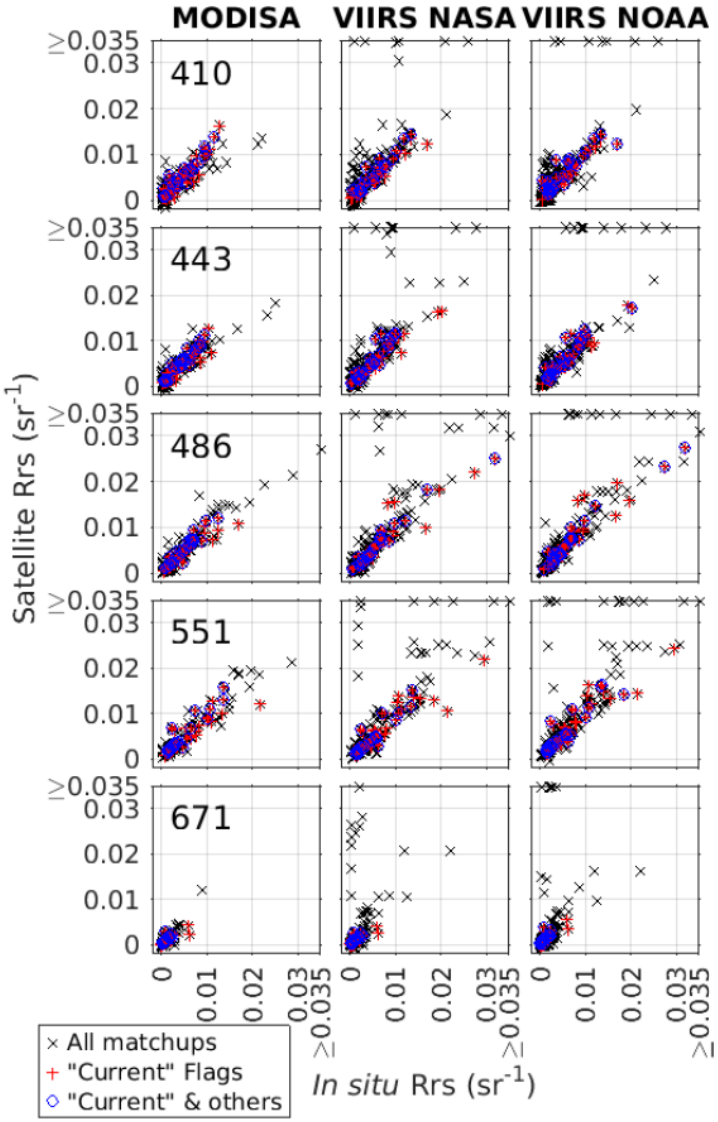
Figure 1: Map of sample locations, concentrated in the Gulf of Mexico and South Atlantic Bight, grouped according to matchup(s) with satellite data. These data were collected from 53 separate cruises of lengths from 1 to 35 days in 2012-2017. Enlargements shown for three regions (1. Florida Panhandle estuaries, 2. Florida Big Bend region, and 3. Tampa Bay) with high sample density (1-3 all have same spatial scale).



From Barnes et al. (2019, RSE)

Validation using data collected in North America

(most data collection supported by NOAA VIIRS cal/val program)



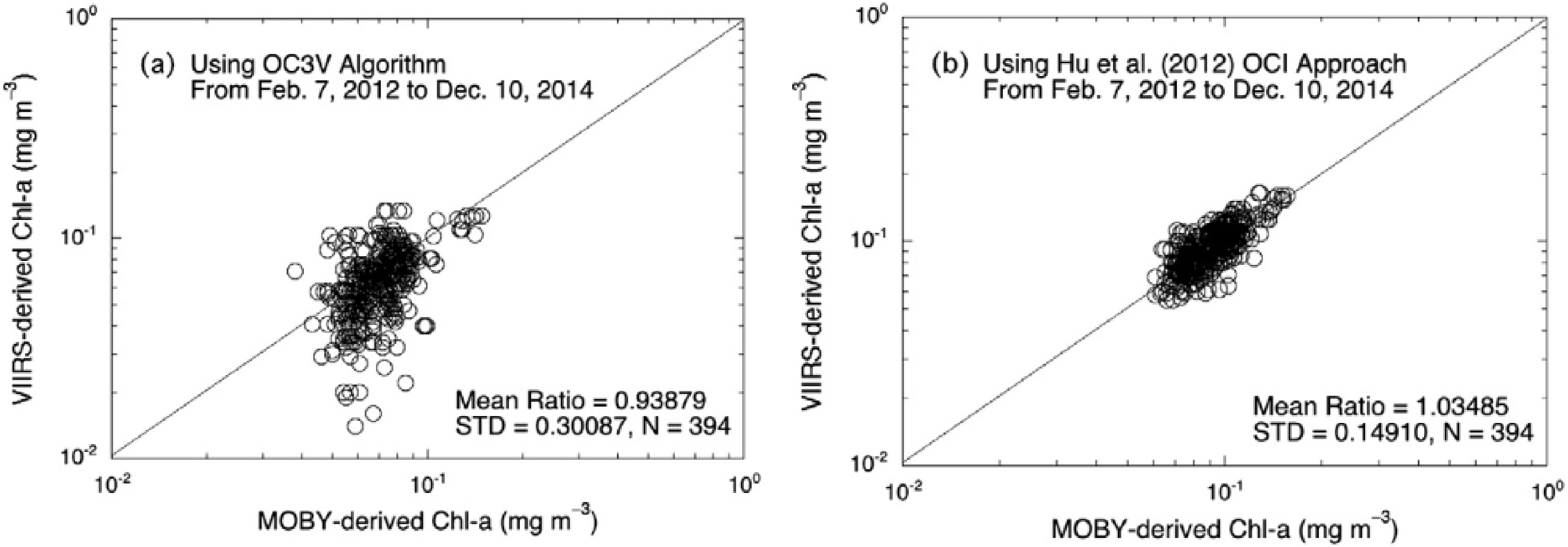
Matchup Statistics

	VIIRS NASA					VIIRS NOAA				
	“Current” Mask, CV < 0.2, +/- 2 h					“Current” Mask, CV < 0.2, +/- 2 h				
Band (nm)	410	443	486	551	671	410	443	486	551	671
UPD (%)	24	24	18	21	26	30	22	18	23	39
MRD (%)	26	25	8	2	9	45	23	12	16	64
RMSE	0.0013	0.0016	0.0017	0.0015	0.0005	0.0018	0.0016	0.0017	0.0019	0.0006
MAPD (%)	34	32	20	23	31	53	31	21	28	71
MR	0.91	0.88	0.98	1.08	1.04	0.87	0.91	0.95	0.96	0.84
R^2	0.9	0.86	0.94	0.86	0.63	0.82	0.86	0.96	0.84	0.58
β_0	0.0006	0.0002	0.0011	0.0002	0.0003	0.0015	0.0011	0.0012	0.0007	0.0002
β_1	1.01	1.13	0.83	0.94	0.75	0.85	0.9	0.85	0.93	1.02
Max (sr^{-1})	0.013	0.01	0.032	0.014	0.003	0.017	0.02	0.032	0.019	0.002
N	25	27	25	23	21	32	30	27	27	24

From Barnes et al. (2019, RSE)

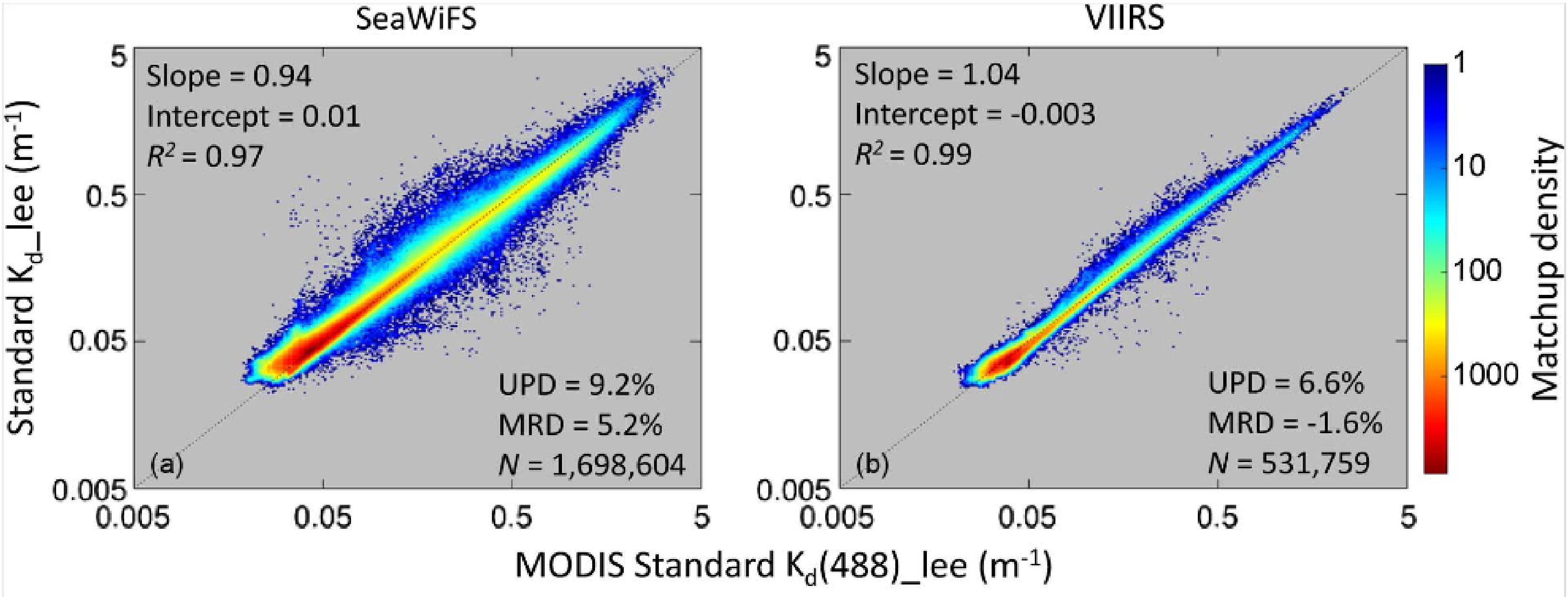
Validation using MOBY measurements

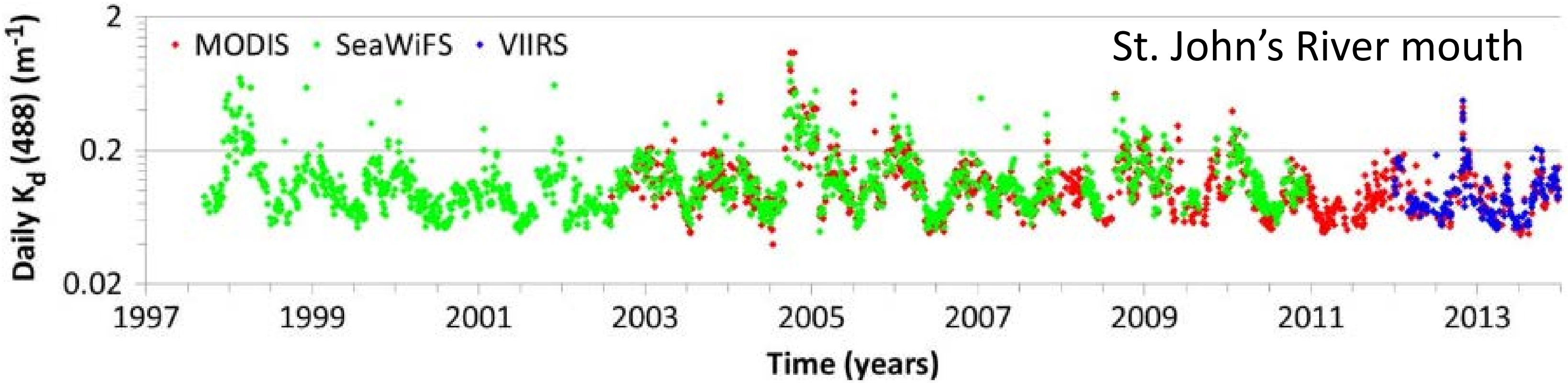
From Wang and Son (2016, RSE)



Validation using cross-sensor comparison

From Barnes and Hu (2015, IEEE TGRS)





1. Data quality: Summary

- Considering that ~40% of the match-ups pairs were collected in shallow (z=3-8m), optically complex coastal waters with variable bottom types, the overall agreement between field and satellite $R_{rs}(\lambda)$ is impressive! [e.g., Uncertainties in Rrs551 ~ 20%, only slightly higher than ~15% from previous reports for less complex waters]
- VIIRS performance is comparable to MODISA performance, from either field validation or cross-sensor comparison
- Results are variable between NOAA and NASA VIIRS products, with neither proving consistently more accurate even when considering only common pixels. Yet NOAA products showed 5-10% more matchup points when identical QA criteria were used

2. Data quantity: evaluation of global data products

Motivation:

- Once data quality is confirmed, the ultimate advantage of remote sensing is in its data quantity (i.e., spatial and temporal coverage frequency), otherwise ship-based measurement is always better
- Then, how to measure this “data quantity”?
- How are different sensors compared with each other? For the same sensor, what differences can result from different algorithms/processings?

2. Data quantity: evaluation of global data products

Measure of “data quantity”

Daily Percentage Valid Observations (DPVOs)

Cloud-free probability over global oceans:

~30% (King et al., 2013)

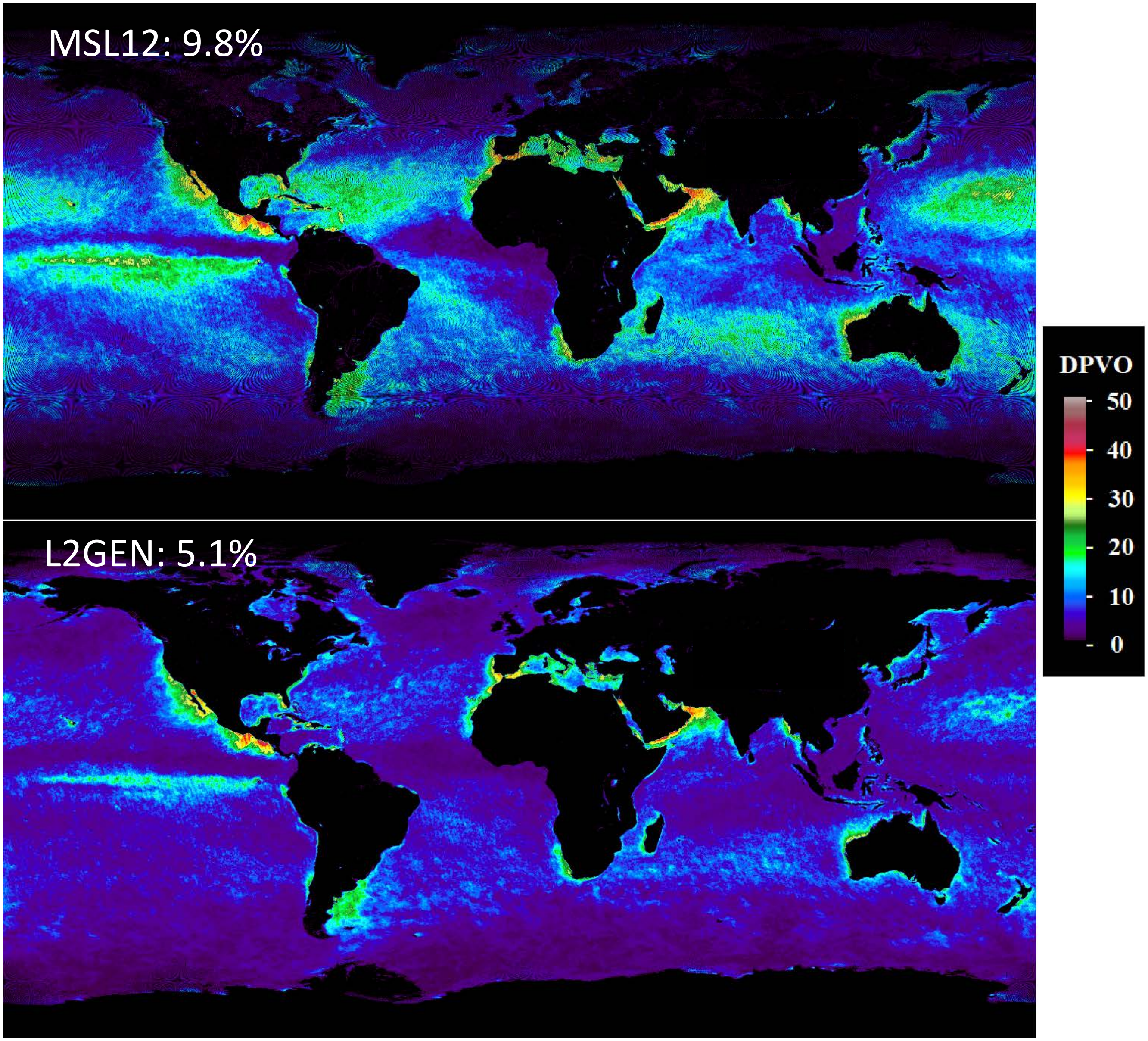
Then, if a sensor can cover global oceans once a day, do we have 30% DPVOs?

Assessment approach:

- # of valid observations (N_v) for each pre-defined 4.6-km grid during each month is stored in global Level-3 bin files
- DPVO = N_v / (4.6 * 4.6 * 30), averaged over global oceans and over many months

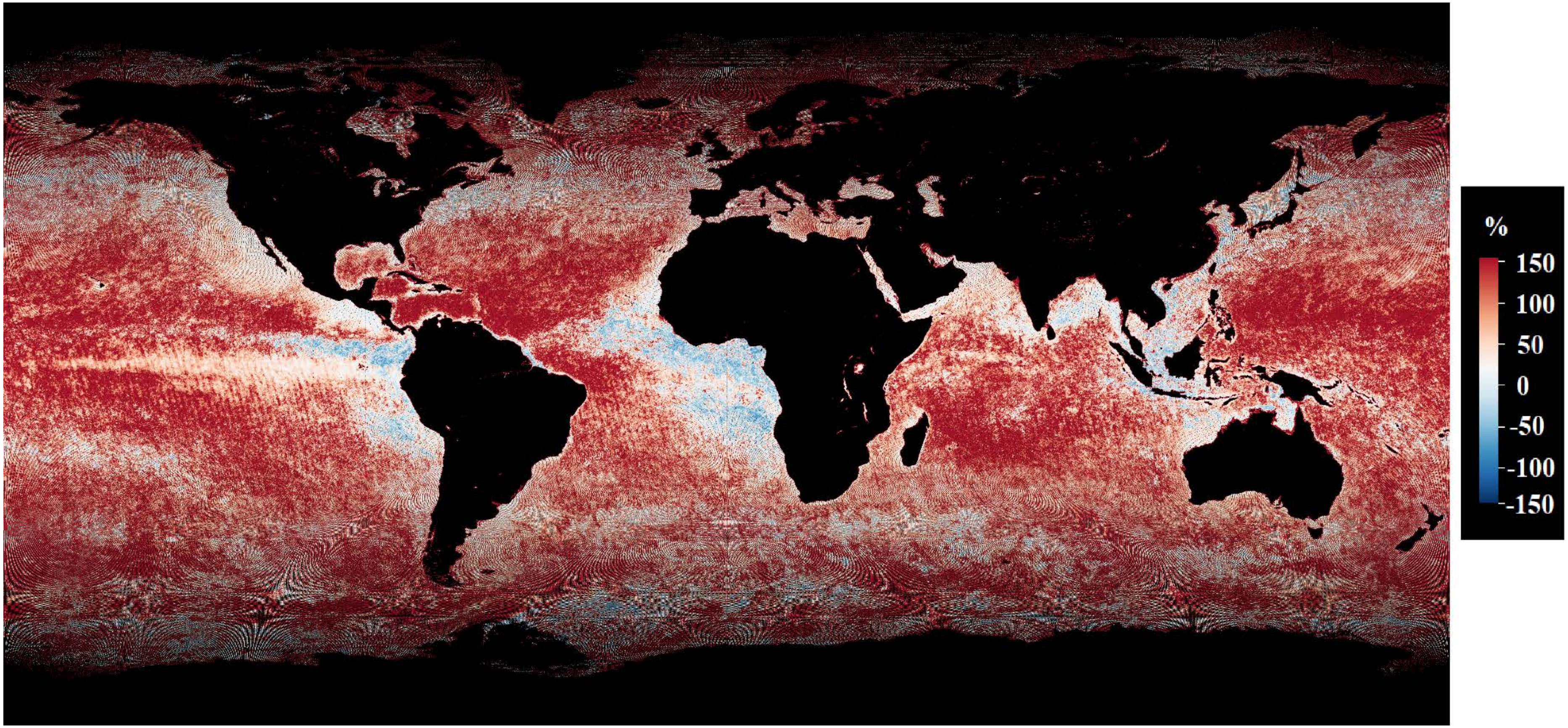
DPVOs of NPP/VIIRS Chl (Nov 2011 – Jul 2018)

NOAA MSL12 versus NASA L2gen



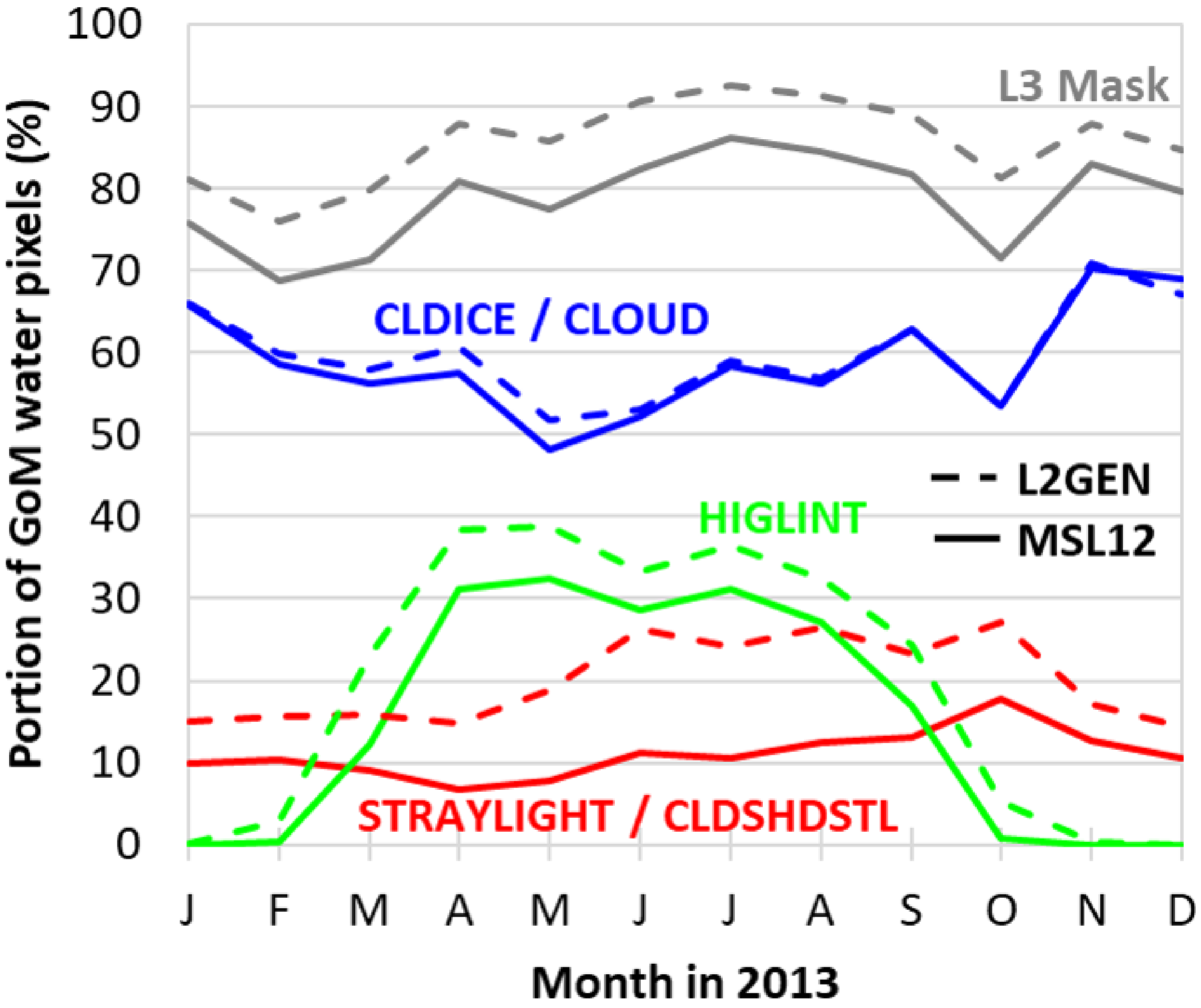
DPVOs of NPP/VIIRS Chl (Nov 2011 – Jul 2018)

(MSL12 – L2gen)/L2gen * 100%



What caused the difference? Gulf of Mexico diagnostics

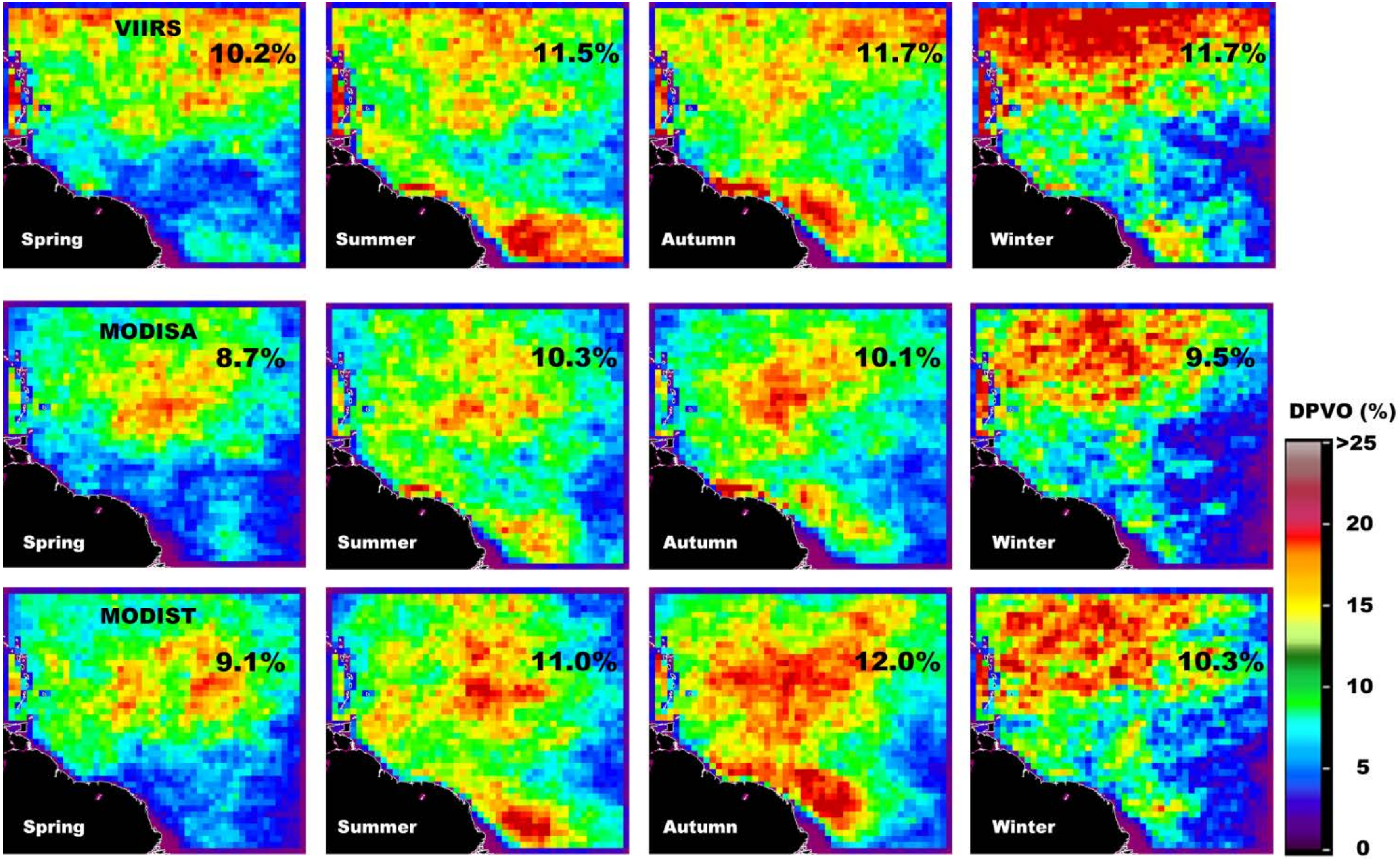
- L2GEN DPVO: 14% (18% in winter, 10% in summer)
- MSL12 DPVO: 21% (25% in winter, 17% in summer)



HIGLINT:
0.01 for MSL12
0.005 for L2GEN

DPVOs of NPP/VIIRS and MODIS AFAI: consistent processing

Daily data from 2006 (Wang and Hu, 2018, IJRS)

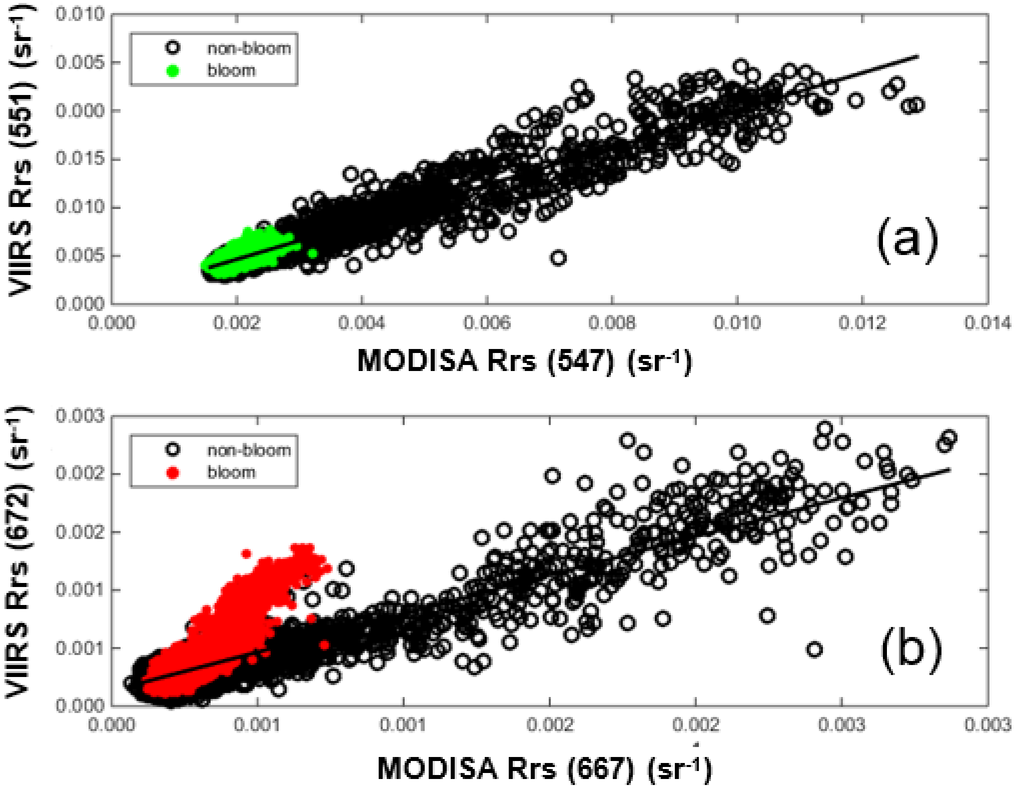


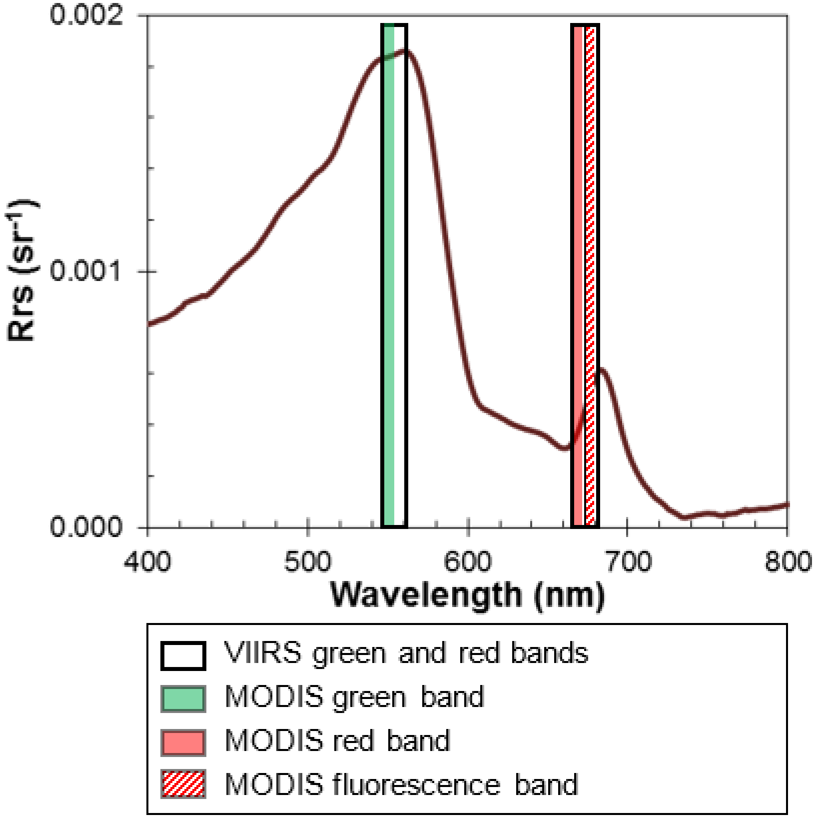
2. Data quantity: Summary

- Even at 30% cloudfree probability, DPVOs of MODIS and VIIRS are way below 30%. Clearly, cloud cover is not the only major reason leading to no valid retrievals
- Straylight and sun glint are two other major reasons
- NOAA MSL12 has a different approach to mask straylight (Jiang and Wang, 2013) than NASA L2gen (7x5 dilation). Together with other differences in processing, MSL12 doubles DPVOs from L2gen on global Chl products
- The increases in quantity have significant implications in tracking blooms and other features as well as in reducing product uncertainties
- Finally, these DPVOs indicate observations at 1-km and daily resolutions. Coverage always increases when data are binned in space and/or time

3. Examples in research and applications: HABs

VIIRS red band (662 – 682 nm) encompasses the MODIS red band (662 – 672) and FLH band (672 – 682), therefore carrying information of algae.

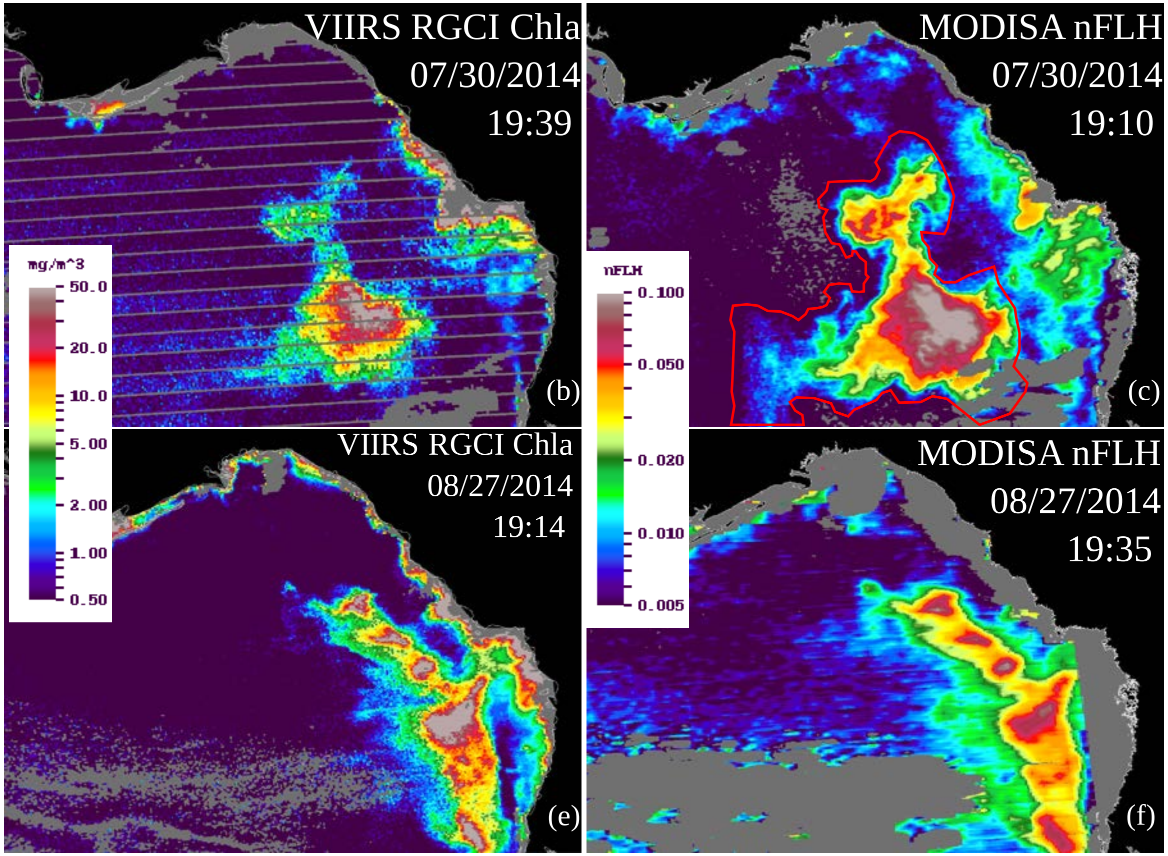




From Qi et al. (2015)

3. Examples in research and applications: HABs

VIIRS RGCI reveals similar patterns as MODIS nFLH (Qi et al., 2015)



3. Examples in research and applications: HABs

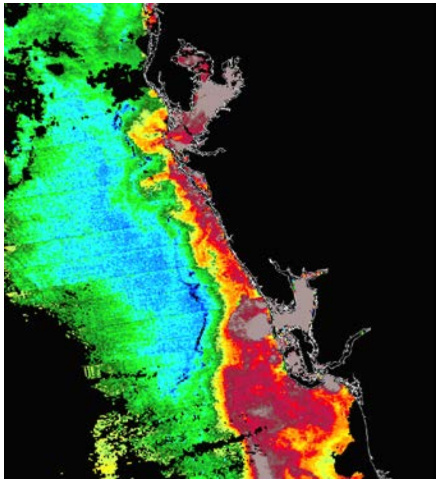
Monitoring of the 2018 red tide in the eastern GOM

VIIRS RGCI can produce products comparable to MODIS nFLH (Qi et al., 2015).

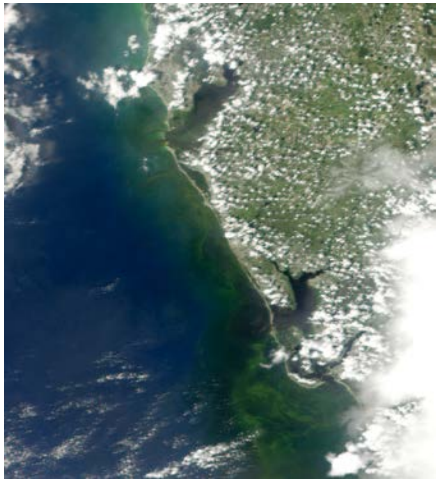
https://optics.marine.usf.edu/projects/iris.html

VIIRS RGB, 8/18/2018

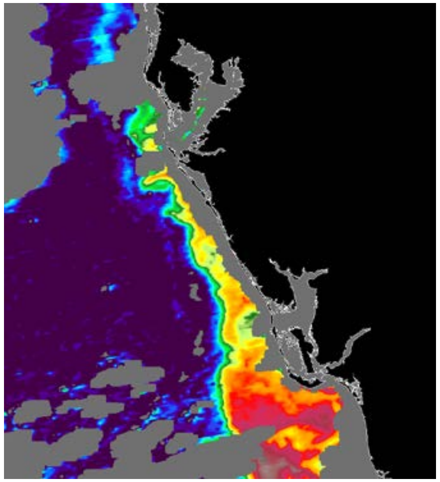
VIIRS RGCI



MODISA RGB



MODIS nFLH



Then, why bother using VIIRS?

3. Examples in research and applications: HABs

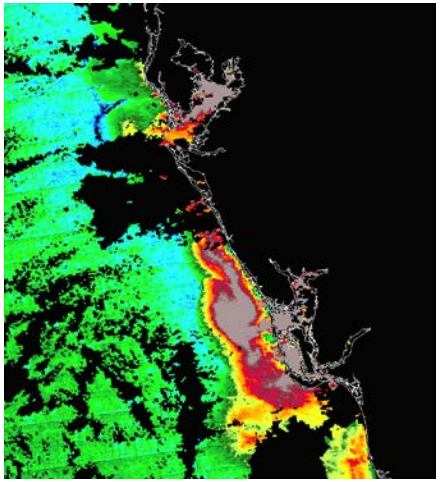
Monitoring of the 2018 red tide in the eastern GOM

VIIRS RGCI can produce products comparable to MODIS nFLH (Qi et al., 2015).

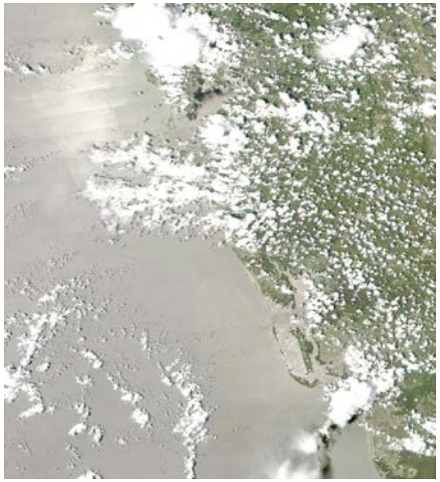
https://optics.marine.usf.edu/projects/iris.html

VIIRS RGB, 8/07/2018

VIIRS RGCI



MODISA RGB



MODIS nFLH

Between 6/1/2018 – 8/22/2018, VIIRS images useful on 39 days (> 50% of coastal regions covered), but MODIS nFLH images useful only on 16 days.

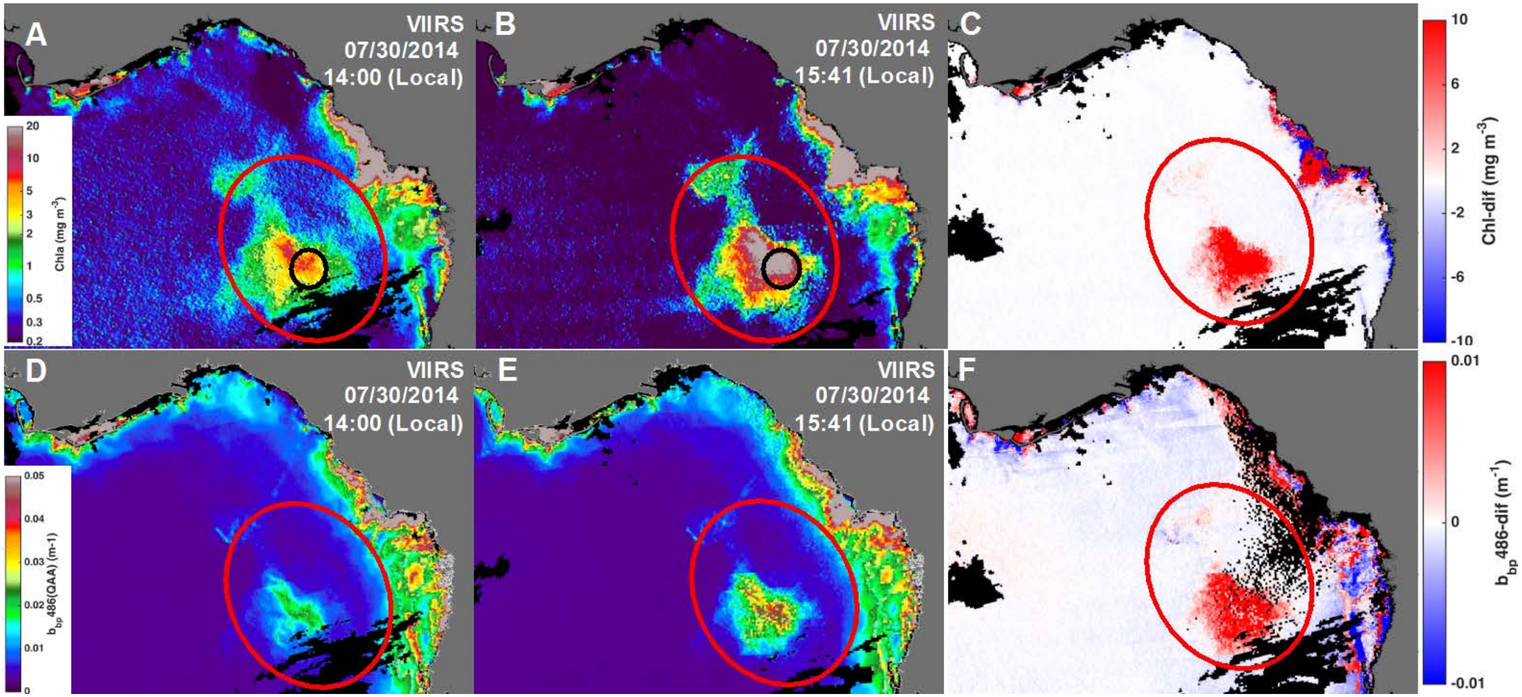
3. Examples in research and applications: HABs

VIIRS captures phytoplankton vertical migration in the NE Gulf of Mexico

14:00 hour 15:41 hour Difference

Chl =>

b_{bp} =>



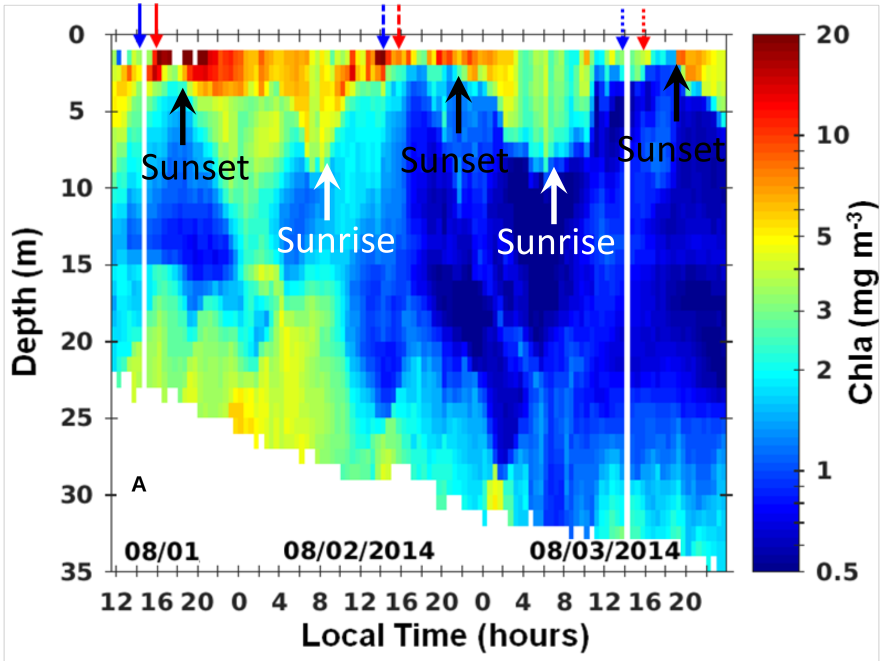
Karenia brevis bloom (red tide), from Qi et al. (2017, Harmful Algae)

3. Examples in research and applications: HABs

VIIRS captures phytoplankton vertical migration in the NE Gulf of Mexico

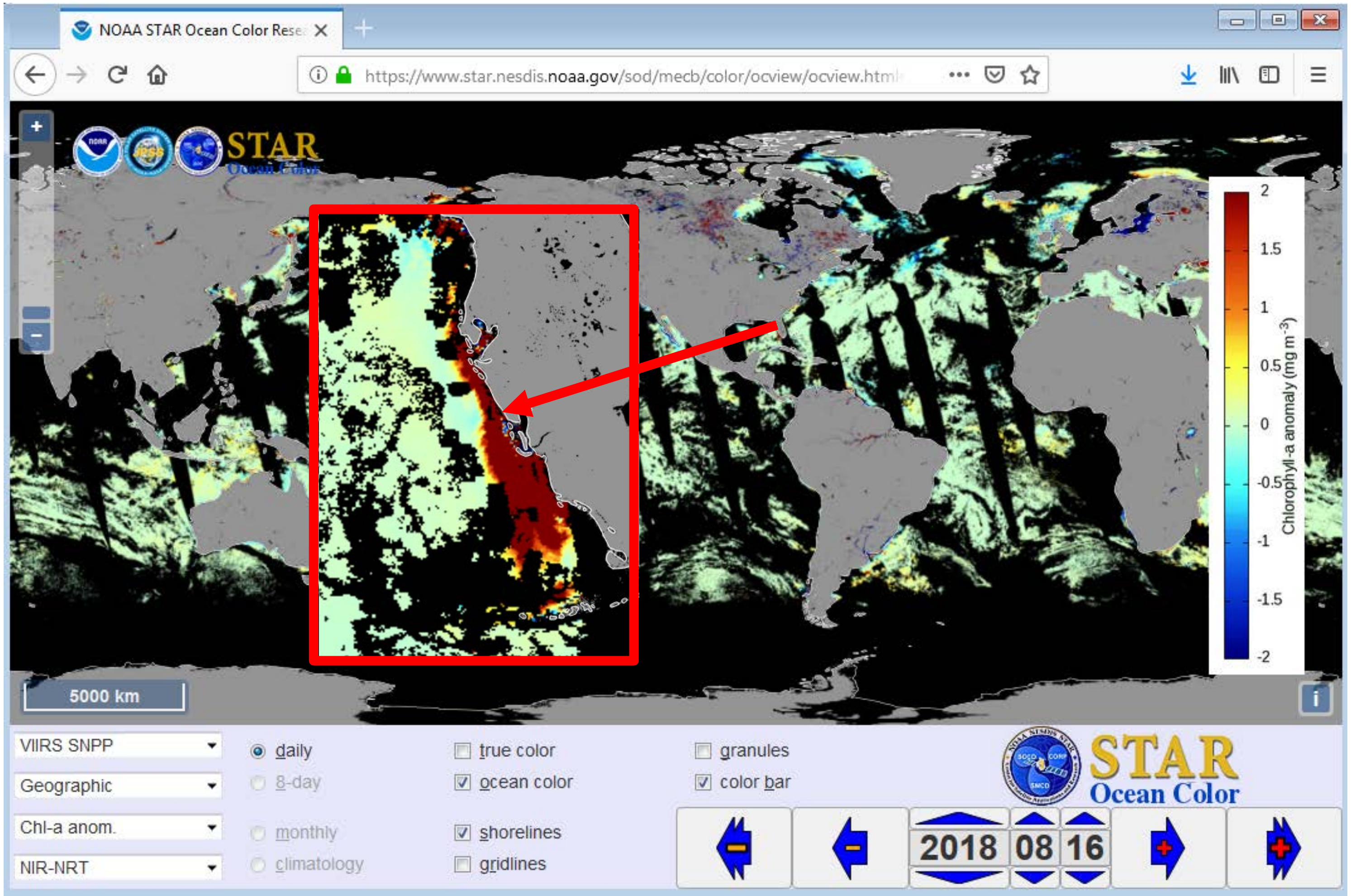
Glider measurement from the same bloom shows thinner surface layer at 15:41 than at 14:00 within a diel cycle of *K. brevis* vertical migration (Hu et al., 2016)

Blue: 14:00; red: 15:41



3. Examples in research and applications: HABs

OCView provides an excellent tool to track blooms in near real time (Mikelsons and Wang, 2018, EOS)



VIIRS SNPP
Geographic
Chl-a anom.
NIR-NRT

● daily
○ 8-day
○ monthly
○ climatology

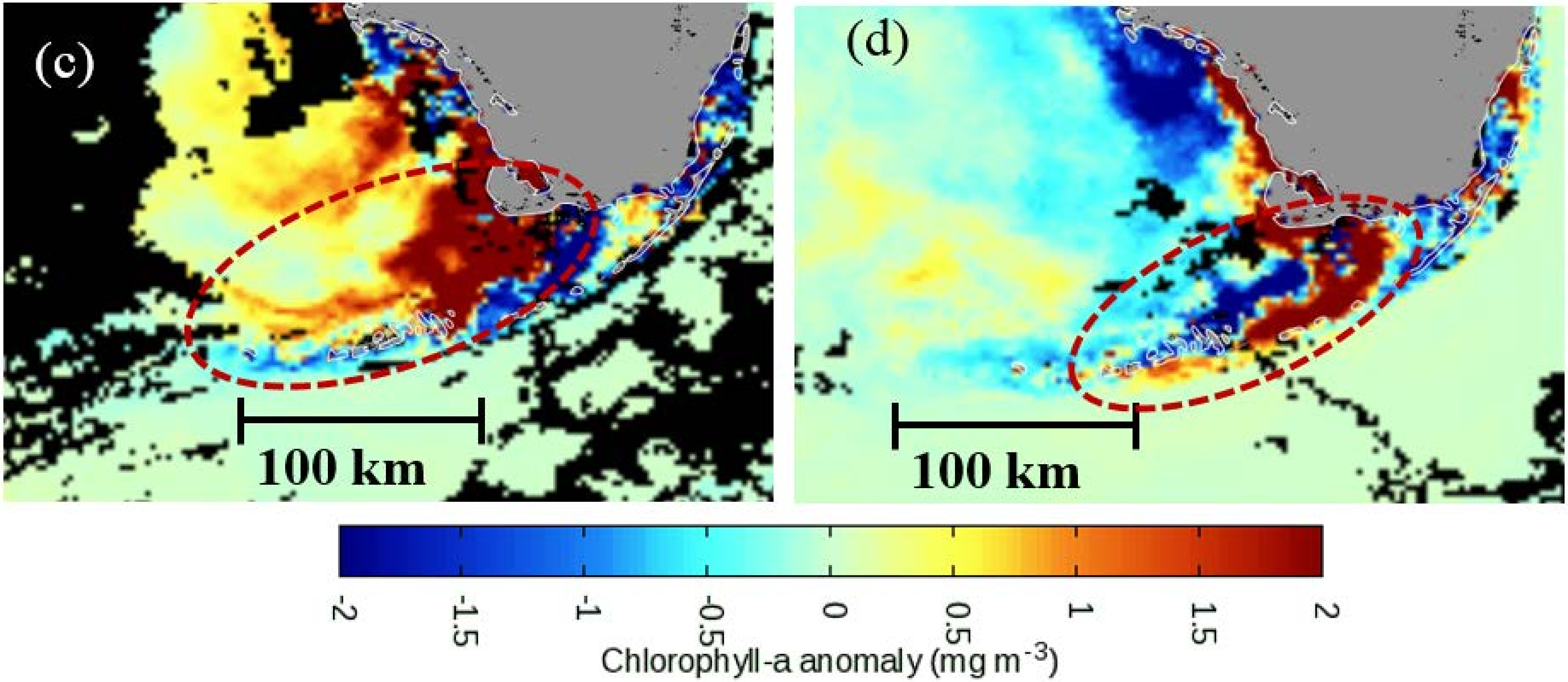
☐ true color
☑ ocean color
☑ shorelines
☐ gridlines

☐ granules
☑ color bar

2018 08 16

3. Examples in research and applications: HABs

OCView Chl anomaly imagery show a *Synecoccocus* bloom (Lapointe et al., Marine Biology, submitted)

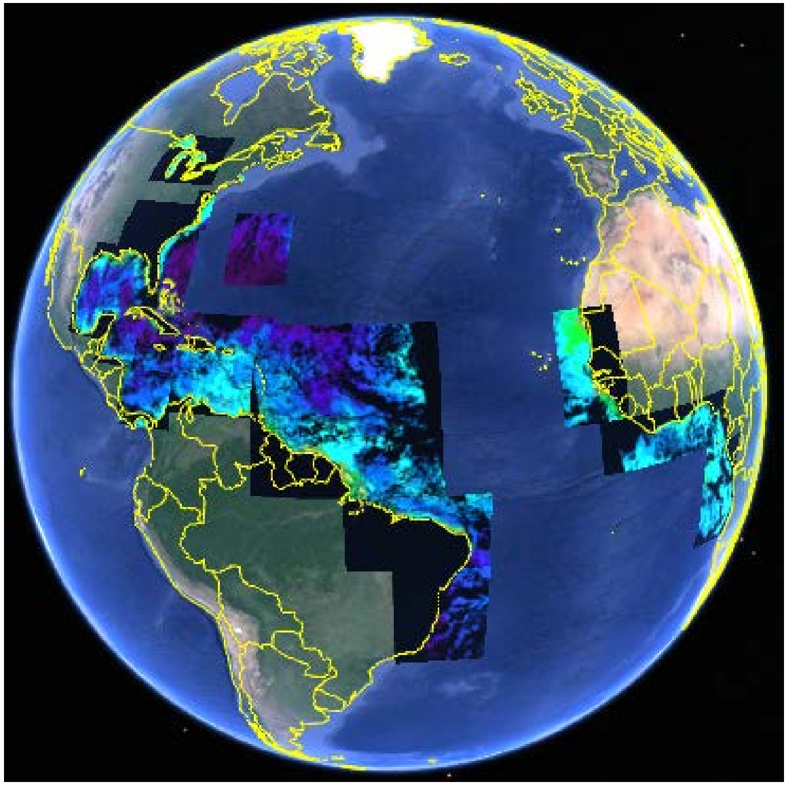


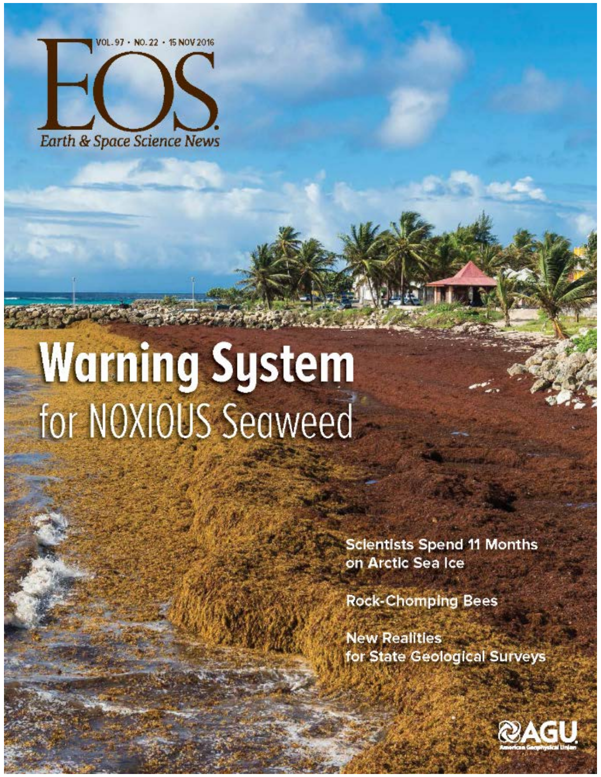
Nov 24, 2013

Jan 27, 2014

3. Examples in research and applications: macroalgae blooms

- Sensors: MODIST, MODISA, VIIRS
- Products: AFAI (1-km), CI (1-km), FAD (5-km)
- Where: https://optics.marine.usf.edu/projects/SaWS.html





VOL. 97 • NO. 22 • 15 NOV 2016

EOS

Earth & Space Science News

Warning System
for NOXIOUS Seaweed

Scientists Spend 11 Months
on Arctic Sea Ice

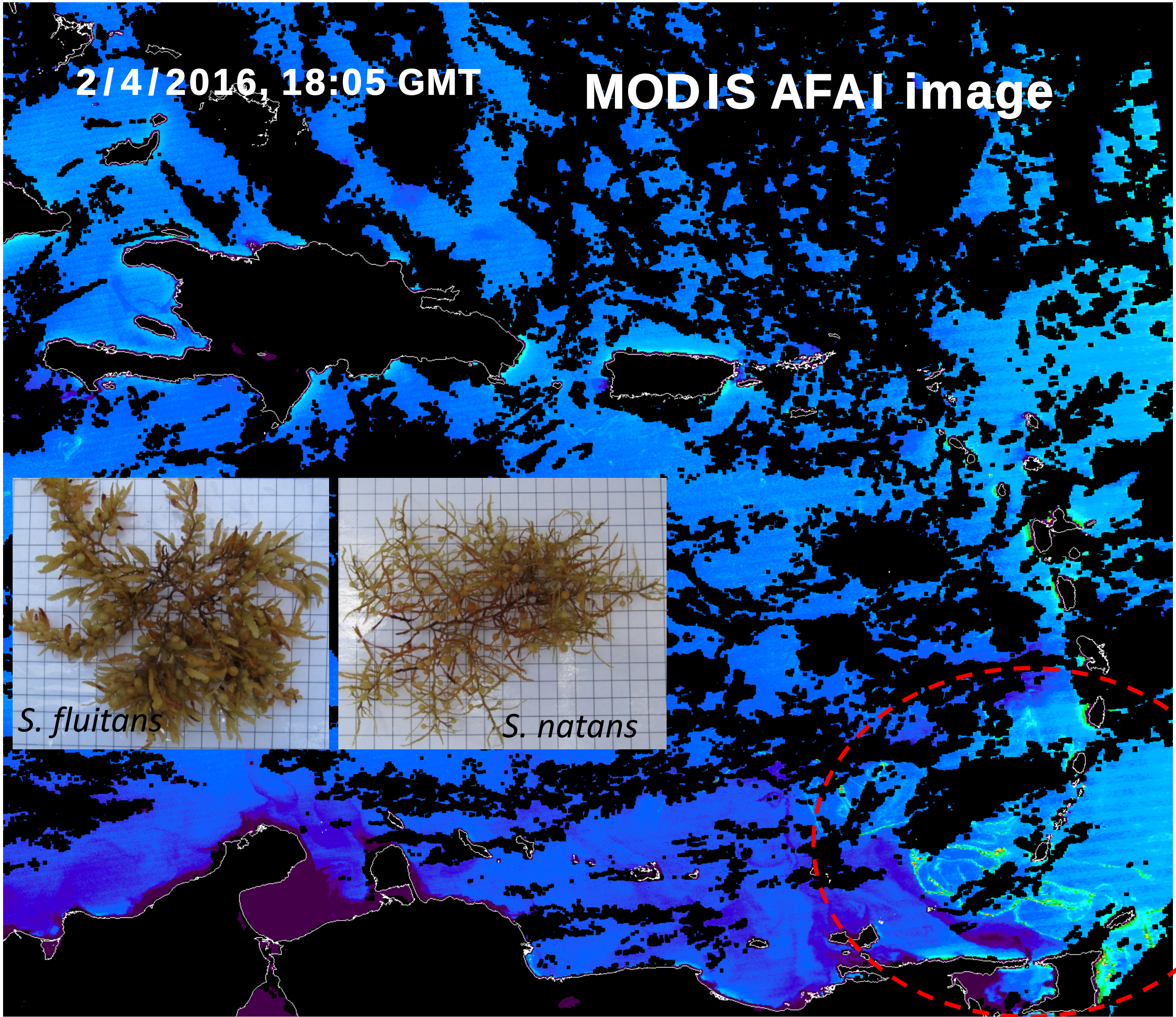
Rock-Chomping Bees

New Realities
for State Geological Surveys

AGU

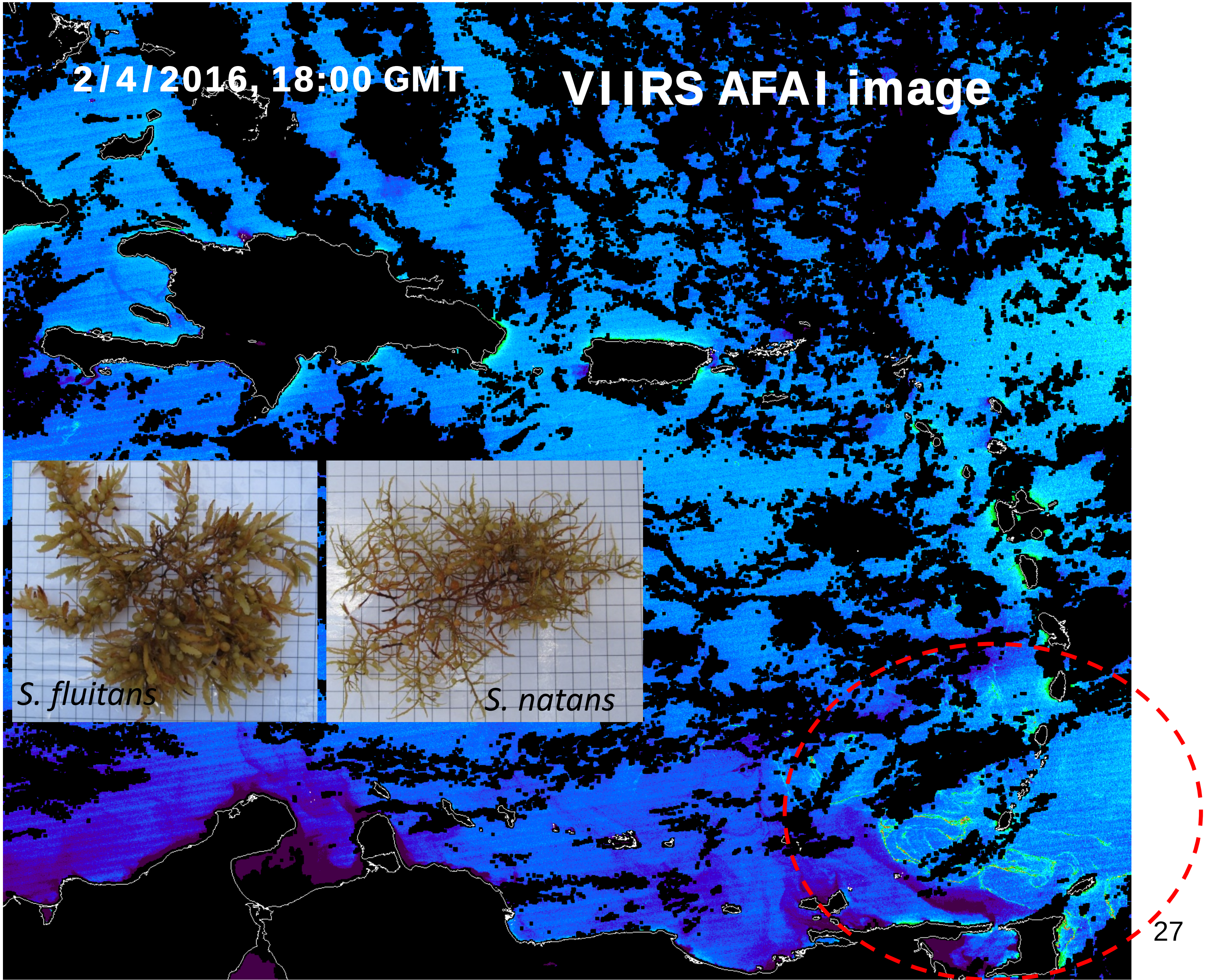
MODIS AFAI image

2/4/2016, 18:05 GMT



VIIRS AFAI image

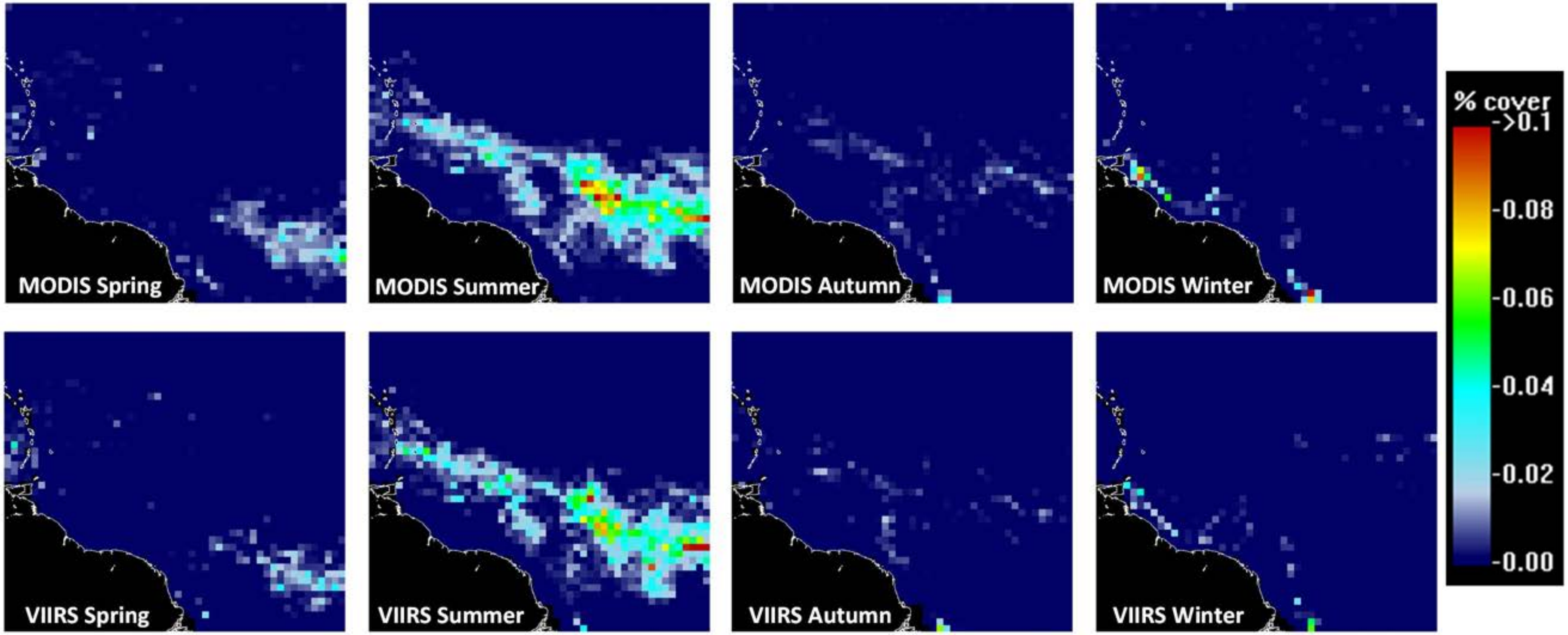
2/4/2016, 18:00 GMT



3. Examples in research and applications: macroalgae blooms

VIIRS continuity in monitoring floating macroalgae in the Atlantic

Color represent mean surface density during 2016 for 0 – 22N, 63 – 38W

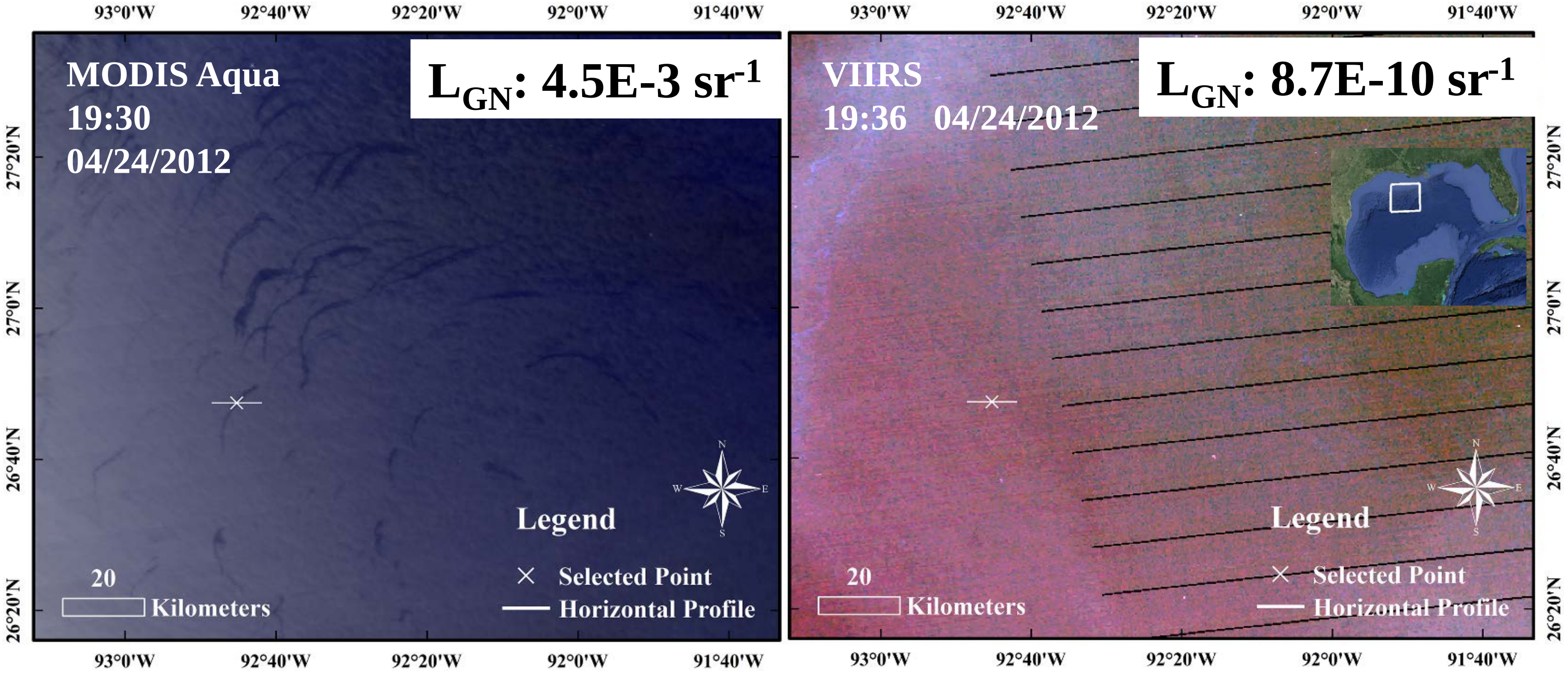


From Wang and Hu (2018). 0.1 on the color scale means 0.1% instead of 10%

3. Examples in research and applications: oil spills

How much sun glint is required to detect thin oil films?

MODIS detects slicks but VIIRS doesn’t

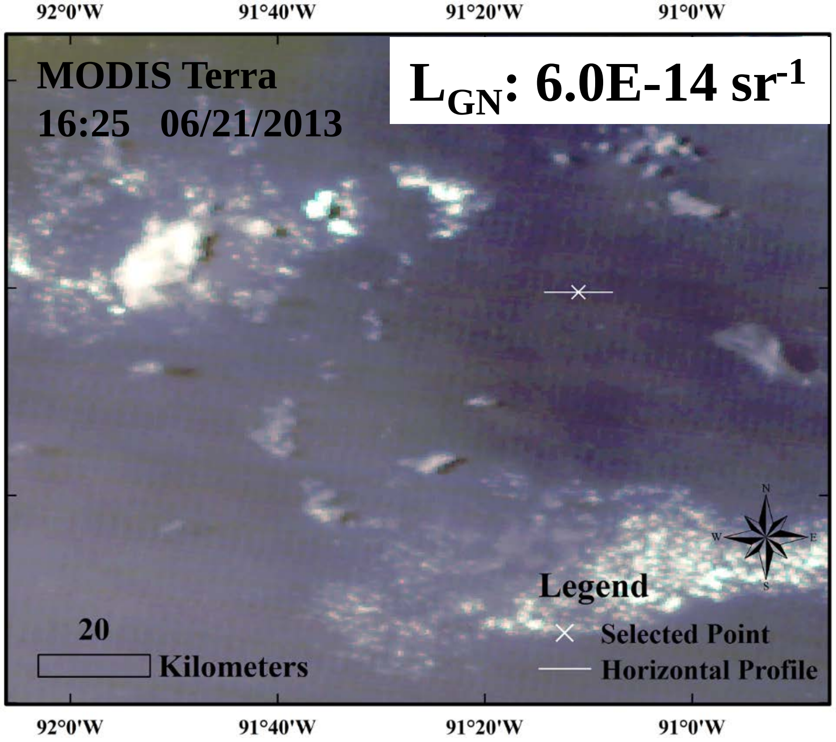


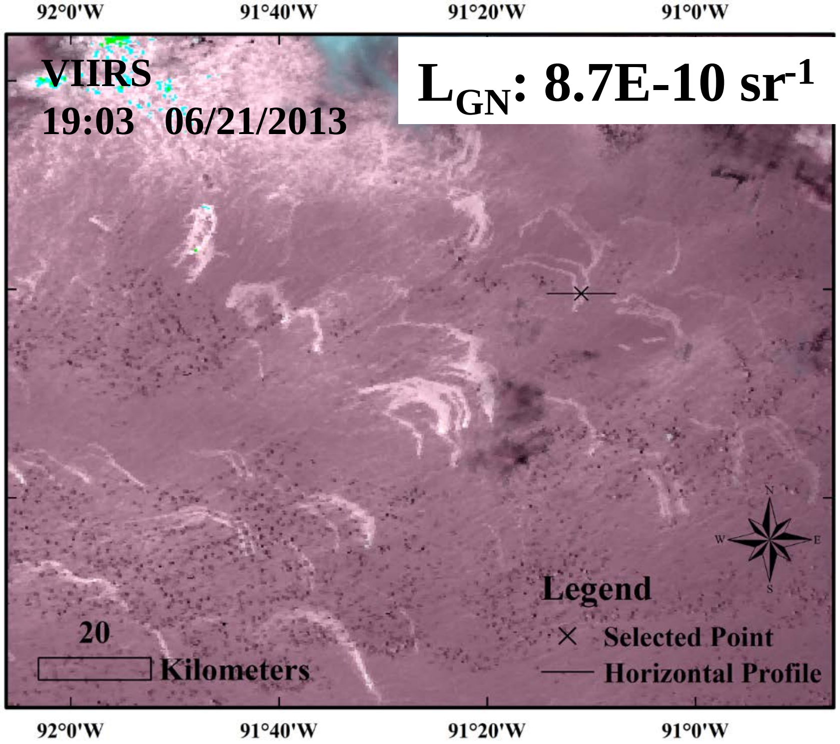
Sun and Hu, 2016

3. Examples in research and applications: oil spills

How much sun glint is required to detect thin oil films?

VIIRS detects slicks but MODIS doesn’t





Sun and Hu, 2016

MODIST requirement: $L_{GN} > 10^{-5}$ sr^{-1}; VIIRS requirement: $L_{GN} > 10^{-6}$ sr^{-1}
VIIRS/MODIS altitudes: 829/705 km; 18% increase in coverage within 60^{o} view angle
VIIRS/MODIS Swaths: 3060/2330 km: 31% increase in all view angle

Summary and Conclusions

- VIIRS/SNPP data products from both NOAA MSL12 and NASA L2GEN show robust performance in northern GoM and N Atlantic waters, similar to MODIS performance
- MSL12 leads to ~10% global DPVO at 1-km resolution, compared to ~5% from L2GEN. This is mostly attributed to their different treatments in stray light and sun glint
- Both data quality and quantity of VIIRS, especially from MSL12, are sufficient for research and applications, as shown in the studies of HABs, macroalgae, and oil spills
- NOAA OCView provides a unique tool for near real-time tracking of anomaly features

Validation using data collected in North America
(most data collection supported by NOAA VIIRS cal/val program)

Temporal distributions

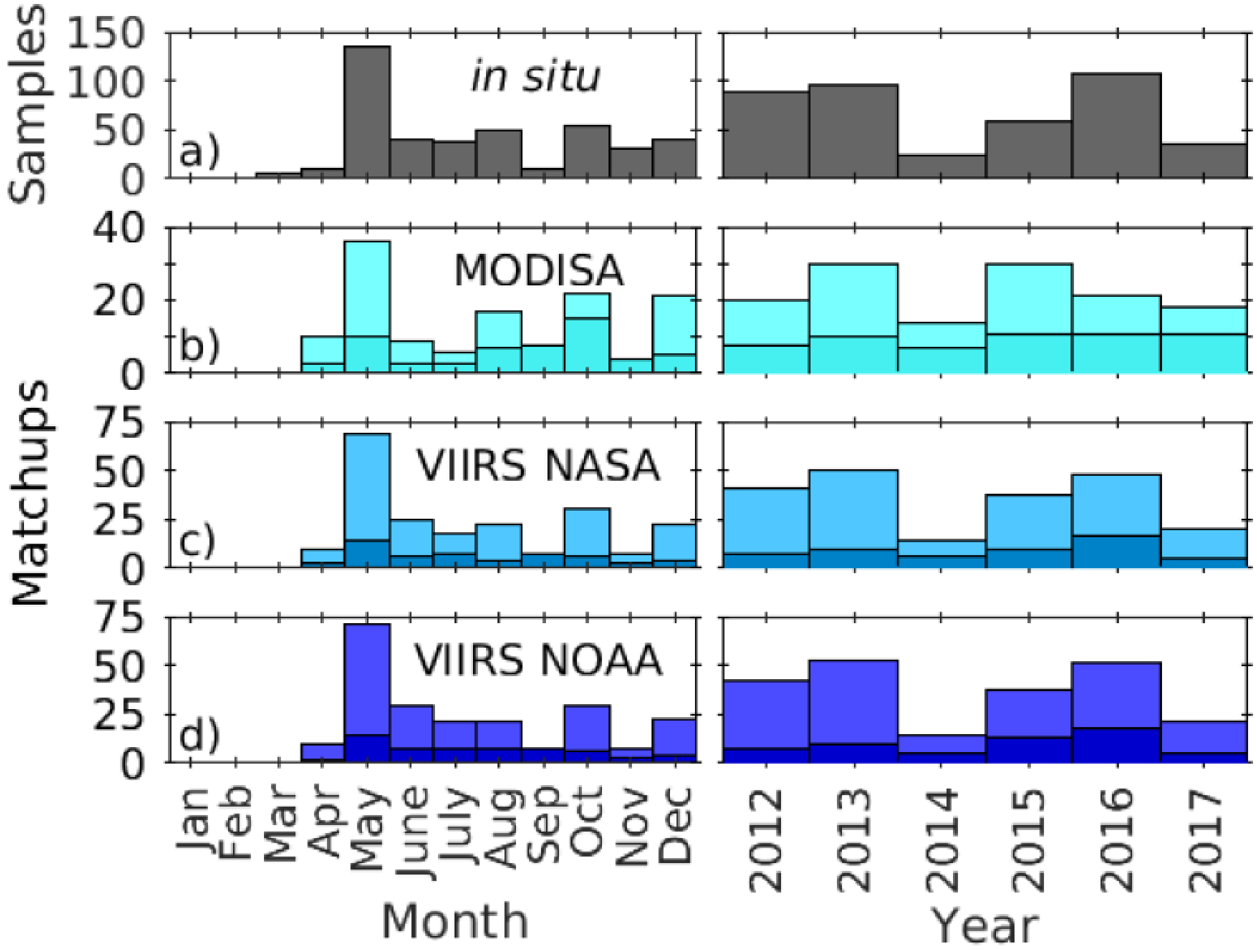


Figure 2: Distribution of (a) *in situ* samples, and (b-d) satellite / *in situ* matchups according to (left column) month and (right column) year. For b-d, lighter color shows any satellite matchups, while darker color excludes matchups identified by the "current" L2 flags (Table 2).

Validation using data collected in North America

(most data collection supported by NOAA VIIRS cal/val program)

Matchup results partitioned by time difference

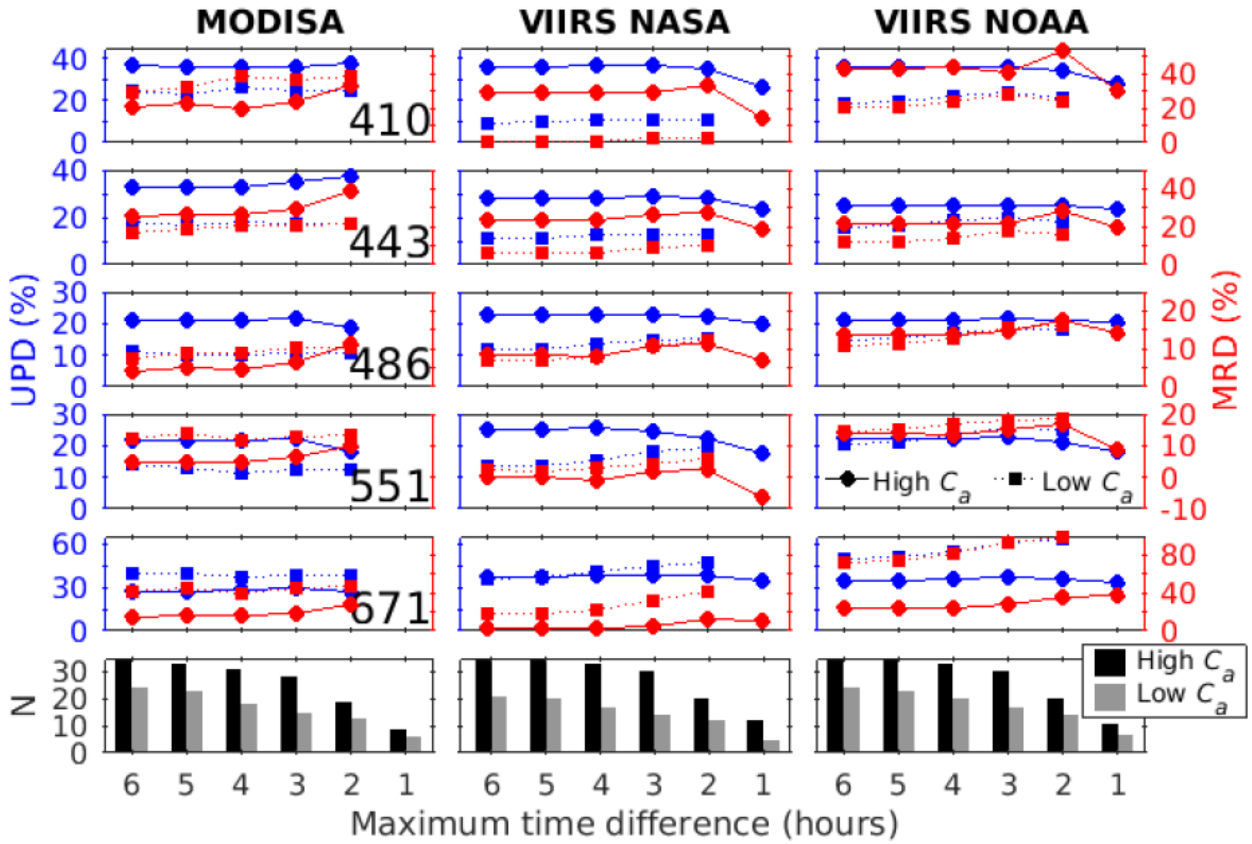
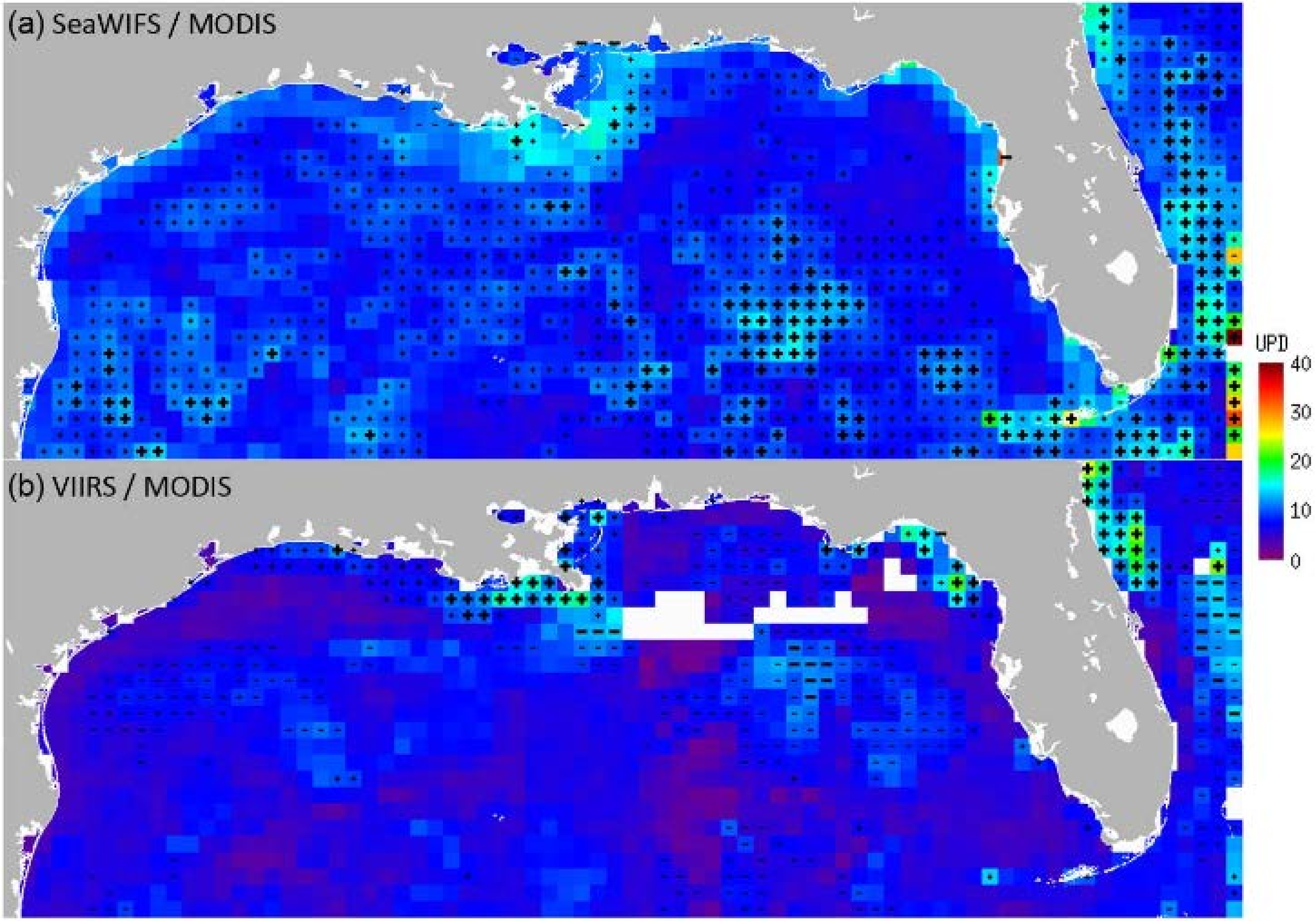


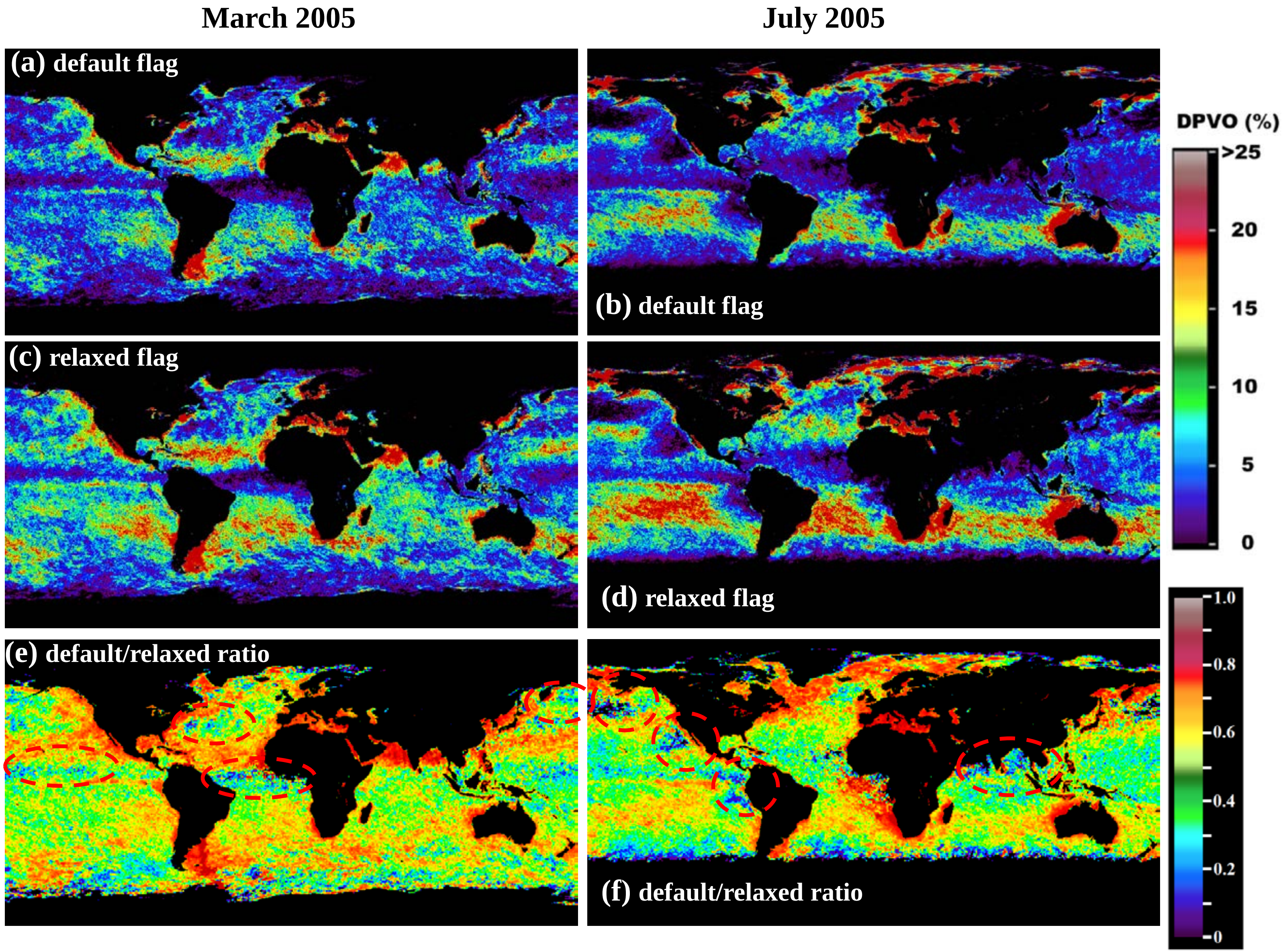
Figure 7: UPD (blue; left axes), MRD (red; right axes), and data quantity (bottom row) for matchup data according to various thresholds of temporal difference between satellite and *in situ* measurements. Data from MODISA (left column), VIIRS NASA (center column), and VIIRS NOAA (right column) are separated by waveband (from top row: 410, 443, 486, 551, and 671 nm), and partitioned into low Ca (dotted lines / squares; water types 1-7) and high Ca (solid lines / circles; water types 8-23). Data partitions with N < 10 are excluded. Unlike Figures 4-6, axis limits are not the same for all wavebands.

Validation using cross-sensor comparison

From Barnes and Hu (2015, IEEE TGRS)



A simple fix using relaxed straylight flag (Hu et al., 2019, JGR)

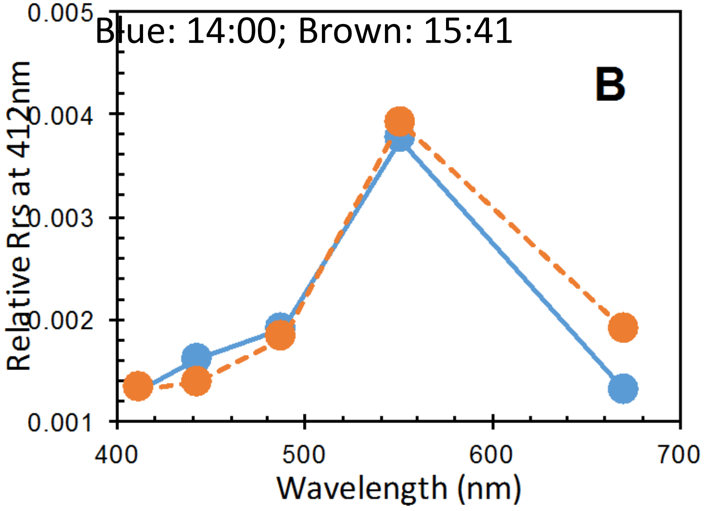


3. Examples in research and applications: HABs

VIIRS captures phytoplankton vertical migration in the NE Gulf of Mexico

Rrs spectral changes in 100 minutes (left) agree with previous field measurement (right)

From VIIRS measurements, July 30, 2014



From Schofield et al. (2006, JGR)