Performance of ocean color retrieval algorithms, verified against in-situ radiometric and sample measurements, show advantages, primarily in complex waters, of algorithms that avoids deep blue bands

Ahmed El-Habashi[1], Michael Ondrusek[2], Vincent Lovko[3] and, Sam Ahmed[1]
[1]NOAA CREST Optical Remote Sensing Laboratory, the City College of New York, CUNY, New York, NY 10031, USA. Email: ahmed.elhabashi@gmail.com, ahmed@ccny.cuny.edu, [2]NOAA /NESDIS 5830 University Research Court, College Park, MD 20740 Email: Michael.Ondrusek@noaa.gov.[3]Mote Marine Laboratories, Sarasota, Florida, United States Email: vlovko@mote.org,

ABSTRACT

We examine the potential for ocean color (OC) retrievals using a neural network (NN) technique recently developed by us to make up for the lack of a 678 nm fluorescence band on VIIRS, previously available on MODIS and important for *Karenia brevis* harmful algal bloom (*KB* HABs) retrievals. NN uses VIIRS Remote Sensing Reflectance (R_{rs}) at 486, 551 and 671 nm to retrieve phytoplankton absorption at 443nm, from which both chlorophyll [*Chla*] concentrations and *KB* HABs can be inferred. NN retrievals are compared with retrievals obtained using other algorithms, including Ocean Color Index (OCI/OCx), Semi-analytical algorithm for both complex and open ocean waters. VIIRS *KB* HABs retrievals in the WFS, using NN and other algorithms, are first compared against all co-incident in-situ cell count measurements available between 2012-16. Next, we compared retrievals obtained for different algorithms using in-situ radiometric R_{rs} measurements against sample measurements, 2017-18, for both the WFS and Atlantic coasts. Next, Retrieval statistics showed (i) the important impact of short term (15-20 minutes) temporal variations and sample depth considerations in complex bloom waters. These limit satellite retrieval accuracies and utility; and (ii) particularly for high chlorophyll bloom waters, better retrieval accuracies were obtained with NN followed by OCI/OCx algorithms. Likely rationales: the longer R_{rs} wavelengths used with NN are less vulnerable (i) to atmospheric correction inadequacies than the deeper blue wavelengths used with other algorithms, and (ii) to spectral interference by CDOM in more complex waters.

Keywords: ocean color, remote sensing reflectance, algorithms development; chlorophyll-*a* retrievals; normalized fluorescence height; neural networks; *Karenia brevis*; harmful algal blooms; West Florida Shelf**.**

1. INTRODUCTION

We have previously described a NN approach [1, 2] for the detection and tracking of *KB* HABs that frequently plague the coasts and beaches of the West Florida Shelf (WFS) using VIIRS satellite data. Effective *KB* HABs detection and tracking approaches are needed for use with the VIIRS satellite. These are needed so that HABs monitoring capabilities can be extend to VIIRS, which previously relied on MODIS-A imagery [3-9] and specifically on (R_{rs}678) the remote sensing reflectance signal at 678nm, the chlorophyll fluorescence wavelength. (R_{rs}678) was used with MODIS-A for the normalized fluorescence height [2-4] and related Red Band Difference (RBD) [3] techniques to effectively help in *KB* HABs retrievals. However, the current VIIRS satellite, unlike its predecessor MODIS-A, does not have a 678 nm channel to detect chlorophyll fluorescence. To overcome the lack of a fluorescence channel on VIIRS, the NN approach bypasses the need for measurements of chlorophyll fluorescence, allowing us to extend *KB* HABs satellite monitoring capabilities in the WFS to VIIRS. Results showed the efficacy of a NN approach for detecting *Karenia* brevis (*KB*) harmful algal blooms (HABs) in the West Florida Shelf (WFS).

2. BACKGROUND

In this work, OC retrievals using NN are comprehensively compared with those using other algorithms, including OCI/OCx and Semi-analytical algorithm for both complex and open ocean

waters. First, VIIRS *KB* HABs retrievals in the WFS, using NN and other algorithms (including OCI/OCx, GIOP, QAA and RGCI), are compared against all co-incident in-situ cell count measurements available over the 2012-16 period. These data sets cover the available data from the start of the VIIRS mission in January 2012 to 2016. In these comparisons, the NN technique was found to exhibit the highest retrieval accuracy statistics followed by OCI/OCx algorithms. The results also highlighted the impact of temporal variations on achievable retrieval accuracies.

Next, we obtained retrievals for NN and the other algorithms using in-situ radiometric R_{rs} measurements as inputs, and compared these retrievals against in-situ sample measurements, for both the WFS and Atlantic coasts. Here, the use of in-situ R_{rs} measurements eliminates the impact of possibly inadequate atmospheric correction procedures that can affect satellite R_{rs} retrievals. Retrieval statistics confirmed (i) the important impact of short term (15-20 minutes) temporal variations and sample depth considerations in complex bloom waters. These limit satellite retrieval accuracies and utility; and (ii) particularly for high chlorophyll bloom waters, better retrieval accuracies were obtained with NN followed by OCI/OCx algorithms. Likely rationales are that: the longer R_{rs} wavelengths used with NN are less vulnerable (i) to atmospheric correction inadequacies than the deeper blue wavelengths used with other algorithms, and (ii) to spectral interference by CDOM, expected in more complex coastal waters. Examination of sequential satellite observations of the WFS also show the importance of short term temporal variabilities and underline their impact on accuracies of retrievals. Finally, field measurements of Sarasota, FL confirm the temporal variability of *KB* HABs and their patchiness in the WFS.

3. APPROACH

For development of our NN algorithm [1, 10-13] a synthetic data set of 20,000 IOPs, was produced based on the NASA Bio-Optical Marine Algorithm Data set, (NOMAD) [16]. These IOPs, whose range and variability is well represented in the literature [17-21] were then used as inputs to a four component bio-optical model [12, 20, 21] which in conjunction with a HydroLight based [22], parameterized forward model, described in Lee 2002 [20] produced 20,000 sets of R_{rs} values at 486, 551 and 671 nm (for VIIRS) and at 488, 555 and 667 nm for MODIS. The NN was trained on 10,000 of these values and tested on the 10,000 remaining subset, as well as on field data to solve the inverse problem [23] of retrieving physical variables, including a_{ph443}, from R_{rs} values at 486, 551 and 671 nm, and at 488, 555 and 667 nm. The algorithm is a standard multiband NN inversion algorithm that takes R_{rs} input at 486, 551 & 671 nm, values used at a longer λ which are not greatly impacted by atmospheric correction. Output of (a_{ph}, a_{dg}, a_{dm} & bb_p) all at 443 which at the peak of a_{ph} and thus exhibit most variation. Detailed descriptions of the NN are given in Refs. [10-13], and [1] gives the necessary parameters for implementation of the NN with a MATLAB tool for obtaining satellite retrievals.

In the following sections, we focus on chlorophyll-*a* concentrations and *KB* HABs retrievals matched against coincident, or near coincident, in-situ measurements in the WFS, using the NN approach and other available direct ocean color retrieval algorithms: (i) Ocean Color indexes (OCI/OCx) product [31, 32]. (ii) Generalized Inherent Optical Property (GIOP) model [26-28] (iii) QAA - Quasi-Analytical Algorithm (QAA) [20, 29, 30] (iv) Red Green Chlorophyll Index (RGCI) [25]. We also examine sources of retrieval's variabilities in the field measurements.

4. RESULTS

3.1 Satellite retrievals and in-situ measurements comparisons for [Chl*a*] and *KB* HABs.

Chlorophyll-a and *KB* HABs retrievals for VIIRS overpass were compared against in-situ measurements using NN and other algorithms. The ultimate test for the viability of *KB* HABs satellite retrieval techniques is their ability to match retrieved values with concurrent *in situ* measurements. However, it is difficult on any one day to find sufficient matchups between satellite observations and concurrent, or near concurrent *in situ* measurements to obtain statistically meaningful results. We therefore extended the study period to look at all available WFS match-ups,

Fig. 1, between VIIRS measurements and *in situ* data at concurrent dates and over the 2012–2016 period for which there was available VIIRS data. We then looked for match ups where the overlap time windows between satellite observations and in-situ measurements were 15 minutes and 100 minutes. The 100 minutes time window conforms to the approximate time between consecutive VIIRS overpasses, and therefore of consecutively retrieved images. These can therefore provide evidence of the temporal variations, observed with *KB* HAB blooms. The shorter 15-minute time window was selected after an investigation that found 15 minutes to be the shortest time window, which at the same time provided enough match up data points for meaningful statistical analysis. This approach is borne out by the results which show much better match ups for the 15 minutes window. This has important implications regarding the validity of satellite observations of *KB* HABs.

When comparisons were being made between retrievals using NN, GIOP, OCI/OCx, QAA and RGCI algorithms, it was found that there were 18 match-ups of available in-situ measurements that satisfy the match-up conditions for satellite observations within a 15 minutes overlap time window. The additional conditions stipulated for match-up were that pixel centers were 0.3 miles or less from the in-situ measurement location. This is an empirical approach to ensure that pixel values could be reasonably assumed to reflect the related in-situ measurements and hence reduce potential impact of patchiness [33] within the pixel (0.7 Km^2 and 1.0 Km^2 corresponds to nadir observation for VIIRS and MODIS respectively). Pixels were also excluded from the match-up comparisons if they had been flagged for any of the following: land, clouds, failure in atmospheric correction, stray light, bad navigation quality, both high and moderate glint, negative Rayleigh corrected radiance, viewing angle greater than 60°, and solar zenith larger than 70°, as well as any pixels which had water leaving radiance spectra with negative values in any one wavelengths. Cell counts sample measurements also had to be at less than 1 meter depth, and at concentrations $\geq 10^4$ cells$\cdot L^{-1}$. It should be noted that [*Chla*] of 1 μg$\cdot L^{-1}$ is taken as ~10^5 cells$\cdot L^{-1}$ [34]. The in-situ cell count data was obtained from the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute (FWC-FWRI). The search for match-up between VIIRS satellite and in-situ observations on the same day and within a 100 minute window of the overpass time showed 93 cases which satisfied the match-up conditions as specified above. To determine the coefficient of determination, R^2, the orthogonal linear regression approach (OR) was used where errors are assumed to exist for both variables. The error (ε) is calculated as the sum of orthogonal distances. OR estimates of X on Y will minimize the orthogonal distance from the observed data points to the regression line [35]. It was found there were 68 valid observations within the 100-minute window between overpass time and the in-situ measurements. The 68 valid observations were reduced to 18 match-ups when the observations window is restricted to 15 minutes, during which the impact of temporal variations is reduced.

Table. 1 shows the results between retrieved [*Chla*] using the NN, GIOP, OCI/OCx, QAA and RGCI algorithms and the in-situ *KB* cell count measurements for the 100 minutes and 15 minute match up time windows are shown. As can be seen, correlations and errors greatly improve for the 15-minute window, compared to the 100-minute window observations. These results, illustrate the impact of observation time windows on retrieval accuracies for all algorithms. It also supports the conclusion, that, at least for these preliminary and somewhat limited data sets, the NN retrievals exhibit the high performances against the in-situ measurements, for both the longer (100-minute) and, more importantly, the shorter (15-minute) overlap time windows. This was observed both in terms of higher correlations and in terms of lower errors against the in-situ measurements.

3.2 Satellite temporal changes on consecutive satellite images

We examine temporal variability here through the use of consecutive satellite images available on VIIRS and MODIS satellites. From the results in 3.1 above, it was seen that reducing the time window between satellite and in situ observations can generally significantly increase the accuracy

between VIIRS retrieved [*Chla*] and in-situ measured *KB* cell counts. These changes can be quite rapid [36]. To explore the potential for detecting HAB bloom changes over relatively short periods from overlapping consecutive satellite overpasses, we have also examined changes in three consecutive overlapping satellite images in Fig. 2. Two of these are from VIIRS, 96 minutes apart and an intermediate one is from MODIS-A, 70 minutes after the first VIIRS image.

We examine retrievals of OCI/OCx [*Chla*] for these three consecutive granules, Fig. 2a, b and c show the retrieved [*Chla*] products from these consecutive VIIRS–MODIS–VIIRS images for the WFS near Sarasota FL on 11/3/2014. The bloom, as delineated by the [*Chla*] color contour in the zoomed images, appears, qualitatively, to increase in concentration and expand in the southwest direction over the 96-minute interval between the consecutive VIIRS–MODIS–VIIRS overlaps. These changes are reflected in the associated zoomed pixel images (Fig. 2) and appear to provide qualitative visual indications of expansion of the bloom and its increasing [*Chla*] concentration in the southwest direction. This appears broadly consistent with the directions of wind and likely current, though there is no specific or quantitative evidence of linkage. Furthermore, no movements of [*Chla*] distribution patterns are discerned that would indicate transport. Nor is there evidence for identifying specific causes for the changes observed, e.g. whether these are due to bloom vertical migration, upwelling/downwelling effects or otherwise. Note bloom free waters (zoomed blue regions in Fig.2, are consistently the same in all 3 images). Consistent bloom areas with no change were also found near the zoomed increasing [*Chla*] concentration regions in Fig. 2 (not shown here).

In conclusion, for this section on consecutive satellite images, it is recognized that additional studies that include comparisons from consecutive VIIRS retrievals of R_{rs} values, as well as comparisons of [*Chla*] retrievals against simultaneous in-situ measurements, would be needed to clarify the nature and magnitude of the changes being observed. However, given the difficulty of carrying out comprehensive calibration measurements of this type, we believe that it is reasonable to conclude that while consecutive overlapping satellite images can provide some evidence of temporal changes in *KB* HABs concentration in the WFS, they are unlikely to provide accurate or reliably useful information on the magnitudes involved.

It should also be noted, that while consecutive overlapping images appear to show temporal changes, there is insufficient evidence from them to attribute the relative contributions of drift, patchiness, upwelling/down welling or a combination of any of these to the causes of the changes. The next section, 3.6 presents the results of recent field measurements of *KB* HABs in the WFS, which much more solidly confirm *KB* HABs temporal variabilities as well as patchiness. They also support conclusions that the significantly improved retrieval accuracies that are obtained with shorter overlap time windows between satellite retrievals and in-situ measurements (section 3.3 above) reflect the impact of temporal variabilities.

3.3 Temporal, intra-pixel and depth effects on field measurements retrievals

The evidence for temporal changes, intra-pixel, depth variations and patchiness associated with blooms in the WFS is further supported more definitively by field measurements made in conjunction with Mote Marine Laboratories on 2017 and 2018 off Lido Key, near Sarasota FL. A transect of measurements were made. Many of these measurements were made at stations subpixel distances apart (generally 300 meters) on an outward leg, and were then repeated for the same stations as closely as possible on a return leg. Radiometers and sample measurements were taken at different depth 0.1, 0.5, and 1 meter. Samples cell concentrations obtained by analysis at Mote Marine Laboratories.

Fig. 3 shows field measured variabilities of *KB* cell counts due to temporal, depth, and intra-pixel effects. *KB* cell counts were converted to equiv. chlorophyll-*a* values in the figure (1 $\mu g \cdot L^{-1} \approx 10^5$ cells$\cdot L^{-1}$ [34]). The measured variabilities of the chlorophyll-*a* values varied from ~1 to 30%, in average. It can be seen, in Fig. 3, that the coefficient of variation (CV) in chlorophyll-*a* values at the same station generally increase with longer time intervals (between outward and return legs). Thus,

the greatest change is for Station 1 from 7.52 x 10^6 cells L^{-1} to 1.552 x 10^6 cells L^{-1} over the 120-minute time interval between the two measurements with a CV=82%. It should be noted, that, as might be expected with these high cell counts, the latter showed an excellent match with simultaneous co-located HPLC [*Chla*] measurements. For the shortest time between measurements, station 2 the change is 0.952 x 10^6 cells L^{-1} to 0.690 x 10^6 cells L^{-1} over the 21-minute interval between measurements with a CV=13%. (It should also be noted that there is also a slight discrepancy in station positions recorded due to drift from measurement start (with GPS initial co-location) to completion and recording. Sampling depths can also result in considerable variability in measured values. The variation for *KB* cell counts values from just below to 0.5 and 1 meters varied in the ranges of 7-32% and 2-24% respectively for 11 stations. For instance, a typical set of measurements of *KB* HABs in the WFS in a 2018 cruise, were for 0.1, 0.5 and 1 meter sampling depths, cell counts/liter were 4.4 x 10^6, 5.1 x 10^6 and 2.4 x 10^6.

These results illustrate both the intra pixel variations that can typically occur (as well as inter pixel variations) and confirm the temporal variations that can be expected. The relative contributions of drift or upwelling/downwelling to the results are not examined here. In general, the consecutive satellite images and the field measurements observations lend support to our underlying thesis that the significantly increased bloom retrieval accuracy that occurs by shortening of the overlap time window between observation and in-situ measurement match-ups from 100 minutes to 15 minutes, is due to temporal changes in the observed bloom. They also serve to underline that magnitudes derived from satellite observations may only be valid for brief periods. To deal with these uncertainties, we are currently examining temporal and spatial averaging possibilities.

3.4 Improving retrieval's performance for satellite and in-situ algorithms.

Better performance of NN VIIRS retrievals against in-situ measurements may be due to relative invulnerability of longer wavelengths used (481, 551, 671 nm) to atmospheric correction inadequacies impacting shorter wavelengths used in other algorithms, particularly for closer in-shore waters. To eliminate atmospheric correction inadequacy concerns, additional credence by examining retrieval of NN and OCI/OCx algorithms using two different blue bands as inputs (443 & 486 nms). Retrieval comparisons were made – for a variety of waters: Open Ocean (Atlantic), and complex coastal *KB* bloom waters. The non-bloom in-situ and satellite Open Ocean measurements are from the NOAA VIIRS Calibration and Validation cruises [37-39] over the 2014-16 period. The coastal bloom in-situ measurements are from our 2017 and 2018 WFS field campaigns. Both datasets are included in the comparisons. In both cases in-situ and satellite R_{rs} values and measured samples were used.

As can be seen, in Fig. 4 the NN and OCI/OCx retrievals give higher retrieval accuracies, with inputs 486 nm instead of 443 nm for both in-situ and satellite retrievals. Again, this tends to confirm that the longer wavelengths used as inputs to the retrieval algorithms are advantageous, whether due to lesser vulnerability to inadequate atmospheric correction, and/or to lesser spectral interference from CDOM, particularly in complex waters.

Retrieval performance for both algorithms is also examined by comparing their retrieval statistics at varying distances from the shoreline (color code in figure 4 for the NN). The coastal water regions located 2 to 16 miles from shoreline, the NN algorithm exhibited the lowest retrieval errors and RMSE of ~ 1.3. For the Open Ocean regions both algorithms exhibited similar errors with RMSE ~ 1.4. The regression lines seem to improve in OCI/OCx for the low chlorophyll-a retrieved values and in NN for the high chlorophyll-a retrieved values. In general, for the more complex WFS waters with *KB* HABs, the NN algorithm performs with higher accuracy, closely followed by OCI/OCx. This would tend to support the notion that higher accuracies of NN retrievals may be at least partially due to the use of longer R_{rs} input wavelengths for NN, than the deeper blue wavelengths used with other algorithms. Thereby lessening the impact that inadequate atmospheric corrections associated with

deeper blue wavelengths will affect retrieval accuracies, as well as possible reduced spectral interference by the existence CDOM (expected in more complex waters).

5. DISCUSSION AND CONCLUSIONS

In the work reported here, NN algorithms using R_{rs} values from the 486, 551 and 671 nm VIIRS bands are used to retrieve [*Chla*] values and *KB* HABs (Example of the NN *KB* HABs retrieval process is illustrated in Fig. 5). Retrievals of *KB* HABs in the WFS from VIIRS overpasses, using NN were then compared with other retrievals using other algorithms including OCI/OCx, GIOP, QAA and RGCI against in-situ sample measurements. These comparisons showed that for the available VIIRS observations, the NN technique appeared to offer good potential for effective retrievals of *KB* HABs cell counts in the WFS, and analysis of retrieval statistics showed the NN technique exhibit higher accuracy, followed by OCI/OCx algorithms. However, more comparisons also showed that when the window between *in situ* observations and satellite overpass measurements was reduced from 100 minutes to 15 minutes, retrieval accuracies greatly improved, with increased correlations and reduced errors. This is now understood to be due to the temporal change in the *KB* HABs scene being observed. That these temporal changes can occur fairly rapidly, was confirmed unambitiously by retrievals of *KB* HABs images in the WFS from consecutive VIIRS –MODIS-VIIRS overpasses, as well as field campaigns of in-situ *KB* HABs measurements in the WFS.

The higher retrieval accuracies of *KB* HABs obtained with NN in the WFS, may be explained by the fact that the NN retrieval algorithm uses longer wavelengths than the deep blue wavelengths used by other retrieval algorithms. The longer wavelengths used are less vulnerable to atmospheric correction inadequacies, as well as to spectral interference from CDOM, both of which are factors likely to be present in the complex *KB* HABs coastal waters of the WFS. It is also important to note that the higher accuracies of NN retrievals may also be inherent because of the training used in the evolution of the NN algorithm. This training was based on a very wide and comprehensive variety of global observations, representing ranges and distributions that are typical for both coastal and oceanic water conditions. To shed more light on these questions, retrieval comparisons were carried out in the complex waters of the WFS and in open ocean waters on the Atlantic coasts. Then, NN and OCI/OCx retrieval algorithms was reinforced by examination of retrievals using both algorithms with input at 443 or 486 nm, which are available in the NASA products. These comparisons were carried out with one set using the in-situ radiometric measurements as inputs, while the other examined satellite retrievals. Again, both sets of retrievals were compared against the in-situ sample measurements. For each set the both algorithms using 486 nm as inputs performed better than the one using 443 nm. Analysis of retrieval statistics again showed higher performance, when using the longer wavelength, especially in complex coastal waters.

In general, it can be concluded that NN retrievals are effective for retrieving *KB* HABs in complex coastal waters such as the WFS. Analysis of retrieval statistics against sample measurements show NN performs better in complex coastal waters than other algorithms. The use of longer R_{rs} input wavelengths may make NN them less vulnerable to atmospheric correction inadequacies and CDOM spectral interference than algorithms, which use deep blue input wavelengths which generally impact retrievals in complex coastal waters. It is important to note that temporal variabilities associated with *KB* HABs impact and limit the utility of satellite retrievals, and suggest the use of other observation platforms, possibly including UAVs.

REFERENCES

[1] A. El-Habashi, I. Ioannou, M. Tomlinson *et al.*, "Satellite Retrievals of Karenia brevis Harmful Algal Blooms in the West Florida Shelf Using Neural Networks and Comparisons with Other Techniques," Remote Sensing, 8(5), 377 (2016).

[2] A. El-Habashi, C. M. Duran, V. Lovko *et al.*, "Satellite retrievals of Karenia brevis harmful algal blooms in the West Florida shelf using neural networks and impacts of temporal variabilities," Journal of Applied Remote Sensing, 11(3), 032408-032408 (2017).

[3] C. Hu, F. E. Muller-Karger, C. J. Taylor *et al.*, "Red tide detection and tracing using MODIS fluorescence data: A regional example in SW Florida coastal waters," Remote Sensing of Environment, 97(3), 311-321 (2005).

[4] R. Amin, J. Zhou, A. Gilerson *et al.*, "Novel optical techniques for detecting and classifying toxic dinoflagellate Karenia brevis blooms using satellite imagery," Optics express, 17(11), 9126-9144 (2009).

[5] M. Tomlinson, T. Wynne, and R. Stumpf, "An evaluation of remote sensing techniques for enhanced detection of the toxic dinoflagellate, Karenia brevis," Remote Sensing of Environment, 113(3), 598-609 (2009).

[6] G. A. Carvalho, P. J. Minnett, V. F. Banzon *et al.*, "Long-term evaluation of three satellite ocean color algorithms for identifying harmful algal blooms (Karenia brevis) along the west coast of Florida: A matchup assessment," Remote sensing of environment, 115(1), 1-18 (2011).

[7] D. Blondeau-Patissier, J. F. Gower, A. G. Dekker *et al.*, "A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans," Progress in oceanography, 123, 123-144 (2014).

[8] C. Hu, B. B. Barnes, L. Qi *et al.*, "A harmful algal bloom of Karenia brevis in the Northeastern Gulf of Mexico as revealed by MODIS and VIIRS: A comparison," Sensors, 15(2), 2873-2887 (2015).

[9] I. M. Soto, J. Cannizzaro, F. E. Muller-Karger *et al.*, "Evaluation and optimization of remote sensing techniques for detection of Karenia brevis blooms on the West Florida Shelf," Remote Sensing of Environment, 170, 239-254 (2015).

[10] I. Ioannou, A. Gilerson, B. Gross *et al.*, "Neural network approach to retrieve the inherent optical properties of the ocean from observations of MODIS," Applied Optics, 50(19), 3168-3186 (2011).

[11] I. Ioannou, A. Gilerson, B. Gross *et al.*, "Deriving ocean color products using neural networks," Remote Sensing of Environment, 134, 78-91 (2013).

[12] I. Ioannou, A. Gilerson, M. Ondrusek *et al.*, "Remote estimation of in water constituents in coastal waters using neural networks." In Proc. of SPIE Vol, vol. 9240, pp. 92400T-1. 2014.

[13] A. El-habashi, and S. Ahmed, "Neural network algorithms for retrieval of harmful algal blooms in the west Florida shelf from VIIRS satellite observations and comparisons with other techniques, without the need for a fluorescence channel." 96380B-96380B-12.

[14] C. Le, and C. Hu, "A hybrid approach to estimate chromophoric dissolved organic matter in turbid estuaries from satellite measurements: A case study for Tampa Bay," Optics express, 21(16), 18849-18871 (2013).

[15] J. P. Cannizzaro, C. Hu, D. C. English *et al.*, "Detection of Karenia brevis blooms on the west Florida shelf using in situ backscattering and fluorescence data," Harmful Algae, 8(6), 898-909 (2009).

[16] P. J. Werdell, and S. W. Bailey, “An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation,” Remote sensing of environment, 98(1), 122-140 (2005).
[17] C. D. Mobley, [Light and water: radiative transfer in natural waters] Academic press, (1994).
[18] R. M. Pope, and E. S. Fry, “Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements,” Applied optics, 36(33), 8710-8723 (1997).
[19] A. M. Ciotti, M. R. Lewis, and J. J. Cullen, “Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient,” Limnology and Oceanography, (2002).
[20] Z. Lee, K. L. Carder, and R. A. Arnone, “Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters,” Applied optics, 41(27), 5755-5772 (2002).
[21] Z. Lee, “Reports of the International Ocean-Colour Coordinating Group: Fundamentals, Tests of Algorithms, and Applications, Report 5,” (2006).
[22] C. D. Mobley, and L. K. Sundman, “Hydrolight 4.2 technical documentation,” Sequoia Scientific, Incorporated, Redmond, WA, 98052, 84 (2001).
[23] F. Aires, C. Prigent, W. Rossow *et al.*, “A new neural network approach including first-guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature and emissivities over land from satellite microwave observations,” (2000).
[24] NASA’s OceanColor Web by the Ocean Biology Processing Group (OBPG) at NASA’s Goddard Space Flight Center. Chlorophyll-a algorithm. (Accessed on 22 November 2016), Available online: https://oceancolor.gsfc.nasa.gov/atbd/chlor_a/.
[25] L. Qi, C. Hu, J. Cannizzaro *et al.*, “VIIRS Observations of a Karenia brevis Bloom in the Northeastern Gulf of Mexico in the Absence of a Fluorescence Band,” IEEE Geoscience and Remote Sensing Letters, 12(11), 2213-2217 (2015).
[26] J. Werdell, “Global Bio-optical Algorithms for Ocean Color Satellite Applications: Inherent Optical Properties Algorithm Workshop at Ocean Optics XIX; Barga, Italy, 3–4 October 2008,” Eos, Transactions American Geophysical Union, 90(1), 4-4 (2009).
[27] P. J. Werdell, B. A. Franz, S. W. Bailey *et al.*, “Generalized ocean color inversion model for retrieving marine inherent optical properties,” Applied optics, 52(10), 2019-2037 (2013).
[28] P. J. Werdell, B. A. Franz, J. T. Lefler *et al.*, “Retrieving marine inherent optical properties from satellites using temperature and salinity-dependent backscattering by seawater,” Optics express, 21(26), 32611-32622 (2013).
[29] Z. P. Lee, QAA Algorithm (Accessed on 22 November 2016), Available online: http://ioccg.org/groups/Software_OCA/QAA_v5.pdf.
[30] Z. Lee, and K. L. Carder, “Absorption spectrum of phytoplankton pigments derived from hyperspectral remote-sensing reflectance,” Remote Sensing of Environment, 89(3), 361-368 (2004).
[31] C. Hu, Z. Lee, and B. Franz, “Chlorophyll aalgorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference,” Journal of Geophysical Research: Oceans, 117(C1), (2012).
[32] J. E. O'Reilly, S. Maritorena, B. G. Mitchell *et al.*, “Ocean color chlorophyll algorithms for SeaWiFS,” Journal of Geophysical Research: Oceans, 103(C11), 24937-24953 (1998).
[33] P. J. Franks, “Spatial patterns in dense algal blooms,” Limnology and Oceanography, 42(5), 1297-1305 (1997).
[34] R. Stumpf, M. Culver, P. Tester *et al.*, “Monitoring Karenia brevis blooms in the Gulf of Mexico using satellite ocean color imagery and other data,” Harmful Algae, 2(2), 147-160 (2003).
[35] L. Leng, T. Zhang, L. Kleinman *et al.*, "Ordinary least square regression, orthogonal regression, geometric mean regression and their applications in aerosol science." 78, 012084.

[36] R. Arnone, R. Vandermeulen, S. Ladner *et al.*, "Diurnal changes in ocean color in coastal waters." 982711-982711-8.
[37] M. Ondrusek, and V. P. Lance, "Report for Dedicated JPSS VIIRS Ocean Color Calibration/Validation Cruise," (2015).
[38] M. Ondrusek, V. P. Lance, M. Wang *et al.*, "Report for dedicated JPSS VIIRS Ocean Color December 2015 Calibration/Validation Cruise," (2016).
[39] M. Ondrusek, V. P. Lance, M. Wang *et al.*, [Report for Dedicated JPSS VIIRS Ocean Color Calibration/Validation Cruise October 2016], (2017).

Abbreviations

The following abbreviations are used in this manuscript:

KB	*Karenia brevis*
HABs	Harmful algal blooms
WFS	West Florida Shelf
VIIRS	Visible Infrared Imaging Radiometer Suite
MODIS-A	Moderate Resolution Imaging Spectroradiometer Aqua
MERIS	MEdium Resolution Imaging Spectrometer
a_{ph}	Absorption coefficient due to phytoplankton particulates (m^{-1})
a_{dg}	Absorption coefficient due to non-phytoplankton particulates and dissolved substances, $a_{dm} + a_g$ (m^{-1})
a_{dm}	Absorption coefficient due to non-phytoplankton particulates (m^{-1})
a_w:	Absorption coefficient due to water (m^{-1})
a_t	Total absorption coefficient, $a_{ph} + a_{dm} + a_g + a_w$ (m^{-1})
b_{bp}	Backscattering coefficient due to particulates (m^{-1})
b_{bw}	Backscattering coefficient due to water (m^{-1})
b_b	Total backscattering coefficient, $b_{bp} + b_{bw}$ (m^{-1})
[*Chla*]	Chlorophyll-*a* concentration ($\mu g \cdot L^{-1}$)
CDOM	Color dissolved organic matter (ppm)
NAP	Non-phytoplankton particulate concentration ($g \cdot m^{-3}$)
AOP	Apparent optical properties
IOP	Inherent Optical properties
RT	Radiative transfer
Rrs	Above-surface remote-sensing reflectance (sr^{-1})
nLw	Normalized water leaving radiance ($W \cdot m^{-2} \cdot \mu m \cdot sr^{-1}$)
MLPNN	Multi Layer perceptron neural network
NN	Neural network
NN [*Chla*]	NN deriving [*Chla*] from *Rrs* as inputs
NOMAD	NASA bio-Optical Marine Algorithm Data set [22].
nFLH	Normalized fluorescence height Algorithm ($W \cdot m^{-2} \cdot \mu m \cdot sr^{-1}$) [55].
RBD	Red Band Difference Algorithm ($W \cdot m^{-2} \cdot \mu m \cdot sr^{-1}$). [5,6]
OC	Ocean Color
OC3	Chlorophyll-*a* concentration ($\mu g \cdot L^{-1}$) derived using VIIRS and MODIS algorithm [40–42].
OCI	Chlorophyll-*a* concentration ($\mu g \cdot L^{-1}$) derived using VIIRS and MODIS algorithm [40].
RGCI	Red Green chlorophyll-*a* Index
GIOP	Generalized Inherent Optical Property
QAA	Quasi-Analytical Algorithm version
R^2	Coefficient of determination. Orthogonal linear regression approach was used.
ε	Error – sum of orthogonal distances.
N	Number of points
μ	Mean value
HPLC	High Performance Liquid Chromatography
MOTE	Mote Marine Laboratory

Figures and Tables

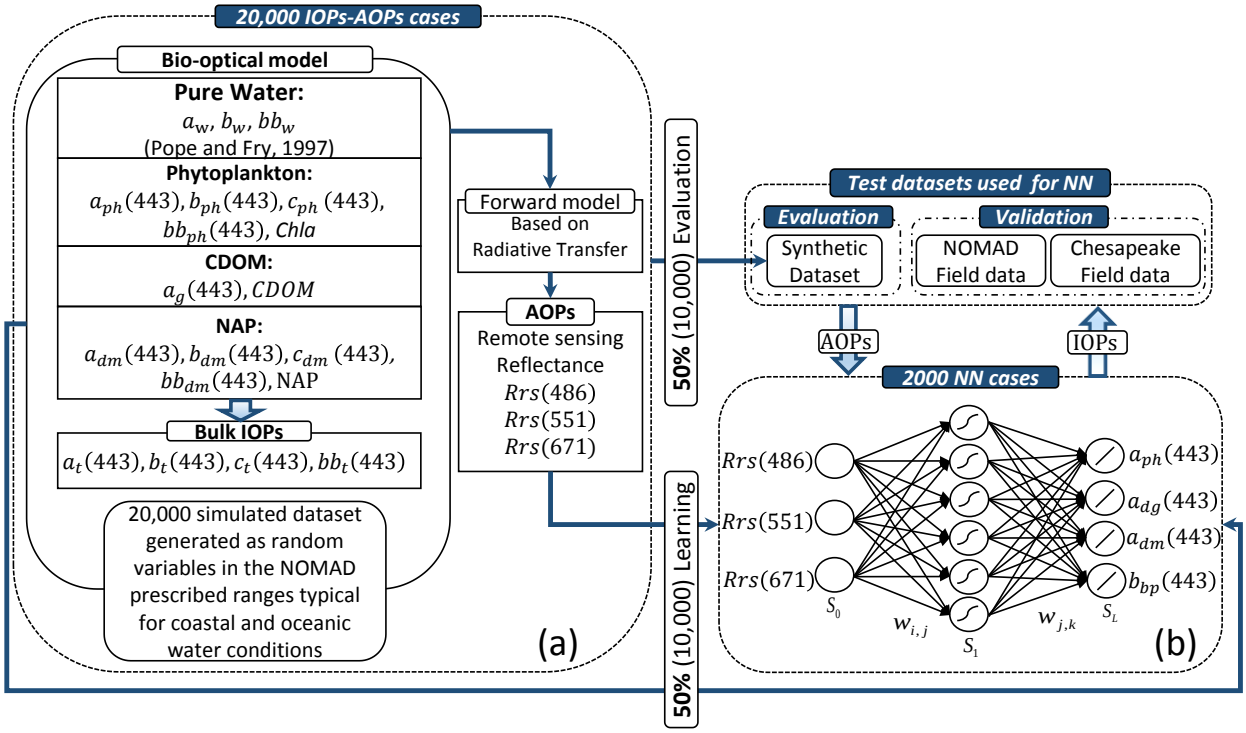

Figure 1 *Shows development of the NN algorithm*. (**a**) Shows the bio-Optical model and forward model simulations; (**b**) Shows the architecture of the NN, one-hidden layer multilayer perceptron (MLP), trained with 10,000 set of R_{rs} and related IOPs.

A synthetic data set of 20,000 IOPs, was produced based on the NASA Bio-Optical Marine Algorithm Data set, [16]. These IOPs, whose range and variability is well represented in the literature [17-21] were then used as inputs to a four component bio-optical model [12, 20, 21] which in conjunction with a HydroLight based [22], parameterized forward model, described in Lee 2002 [20] produced 20,000 sets of *Rrs* values at 486, 551 and 671 nm (for VIIRS) and at 488, 555 and 667 nm for MODIS. The NN was trained on 10,000 of these values and tested on the 10,000 remaining subset, as well as on field data to solve the inverse problem [23] of retrieving ocean color physical parameters, including a_{ph443}, a_g and a_{dm} as well as bb_p,, from *Rrs* values. The NN algorithm uses *Rrs* input at 486, 551 & 671 nm, values used at a longer λ which are not greatly impacted by atmospheric correction. Output of (a_{ph}, a_{ds}, a_d & bo_b) all at 443 which at the peak of a_{pe} and thus exhibit most variation.

Satellite retrieval's statistics of comparison

y-axis ***Retrieved*** *($\mu g \cdot L^{-1}$)*	*x-axis* ***Measured*** *(cells L^{-1})*	R^2	ε	***Slope & Intercept***	***N***
NN	**VIIRS *vs.* in-situ** **100 minutes window**	0.42	5.01	$y = 0.52\,x - 2.04$	68
OCI/OCx		0.34	5.24	$y = 0.43\,x - 1.54$	68
GIOP		0.37	12.64	$y = 1.60\,x - 7.28$	68
RGCI		0.19	15.46	$y = 1.14\,x - 5.39$	68
QAA_{v5}		0.23	15.13	$y = 1.31\,x - 6.09$	68
NN	**VIIRS *vs.* in-situ** **15 minutes window**	0.79	0.42	$y = 0.47\,x - 1.83$	18
OCI/OCx		0.55	0.72	$y = 0.34\,x - 1.13$	18
GIOP		0.45	3.57	$y = 1.20\,x - 5.25$	18
RGCI		0.50	2.28	$y = 0.75\,x - 3.55$	18
QAA_{v5}		0.34	3.82	$y = 0.90\,x - 3.88$	18

Table 1 *Shows the important impact of short term (15-20 minutes) temporal variations and sample depth considerations in complex bloom waters retrievals.* Results are shown for six ocean color retrieval algorithms.

Parameters:

N: Number of data points

R^2: coefficient of determination using the orthogonal linear regression approach (OR).

ε: sum of orthogonal distances from the observed data points to the regression line.

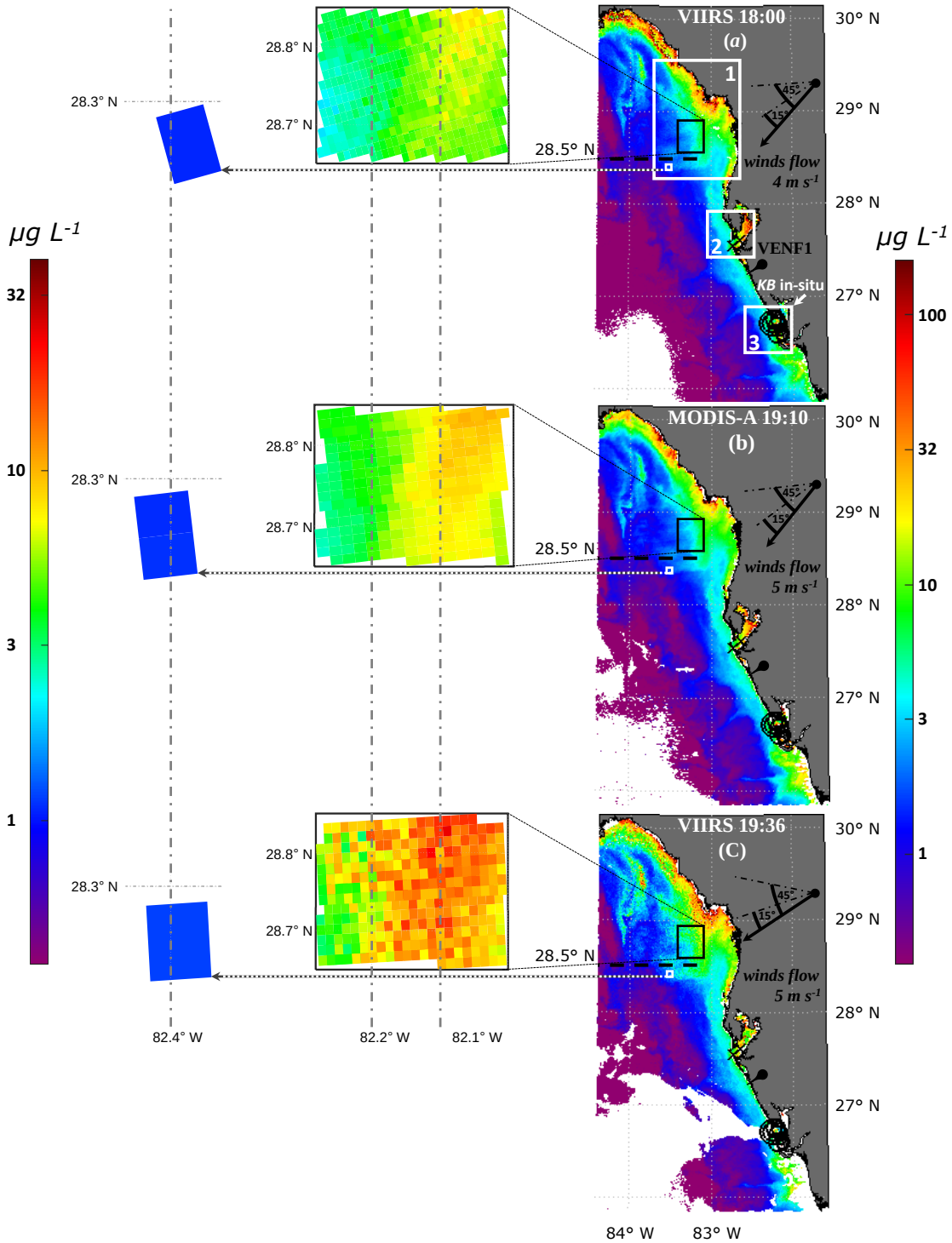

Figure 2 a, b and c *showing 100 minutes changes in bloom for consecutive satellite images of region 1, retrieved chlorophyll-a imagery on Nov. 14th, 2014*. Wind direction information obtained from the National Data Buoy Center website for the C-MAN stations at Venice, FL (Station VENF1). (27°4'21" N 82°27'10" W) gives area wind direction information. Approximate wind and assumed current information is shown overlaid in the retrieval images: Fig. 7.

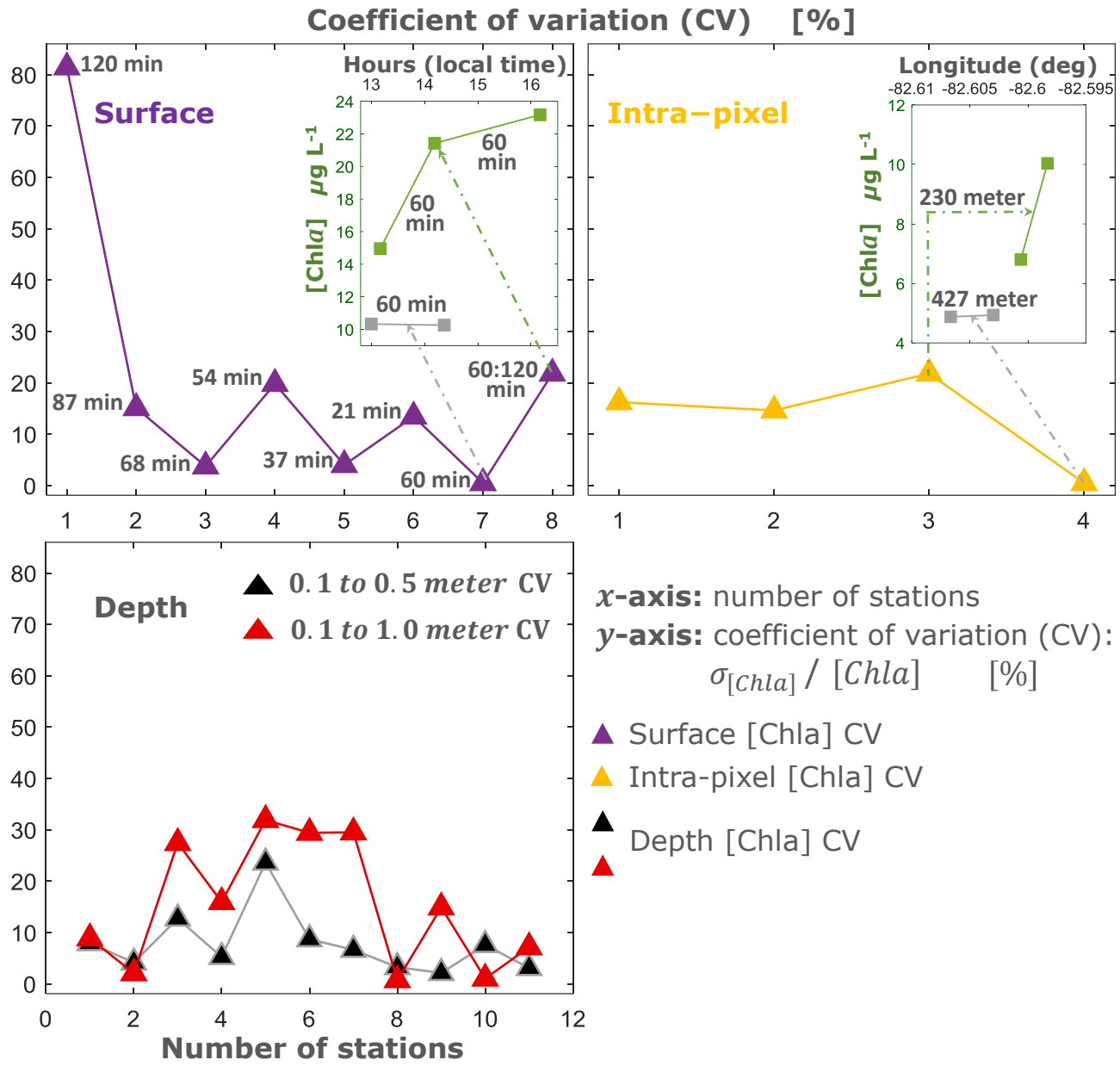

Figure 3 *Shows chlorophyll-a measurements variabilities in coastal bloom waters* (chlorophyll-*a* values shown here are values equiv. to in-situ measured *KB* blooms cell counts). Results show variations ranged from 1 to 30% in average, due to temporal, intra-pixel and depth changes.

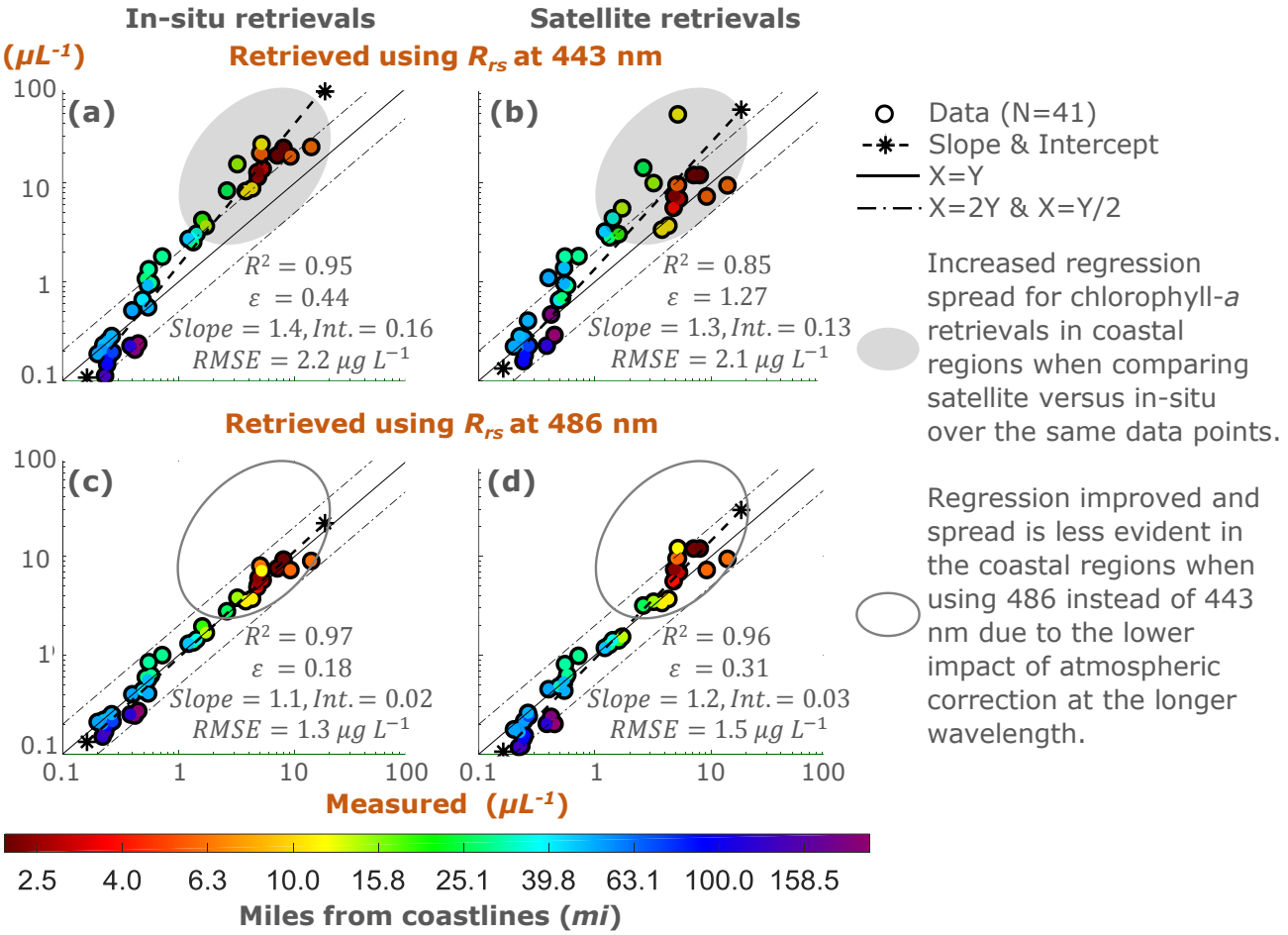

Figure 4 *Shows how the wavelength choice can impact precision and accuracy of chlorophyll-a retrievals.* Color bar shows retrieved [*Chla*] statistics at varying distance from the shoreline.

NN Chlorophyll-*a* retrievals from satellite & in-situ match-ups are compared for the same number of points when using two different blue bands (R_{rs} 443 & 486 nms) as inputs. Results shows improved regression performance in both in-situ and satellite retrieval when using R_{rs} 486 input instead of 443. The use of longer R_{rs} input wavelengths (486nm) make NN retrievals less vulnerable to atmospheric correction inadequacies and CDOM spectral interference than algorithms, which use deep blue input wavelengths (443nm) which generally impact retrievals in complex coastal waters. Similar result is achieved when using the OCI/OCx algorithms. The regression performances of both algorithms are tabulated in the next page for comparison.

Blue band input (nm)	R_{rs} inputs source	Chlorophyll-*a* retrievals using different blue bands (**N=41**)									
		NN algorithm					**OCI/OCx** algorithms				
		R^2	ε	*RMSE*	*Slope*	*Int.*	R^2	ε	*RMSE*	*Slope*	*Int.*
443	**In-situ**	0.95	0.44	2.2	1.4	0.16	0.94	0.57	2.7	1.4	0.28
	Satellite	0.85	1.27	2.1	1.3	0.13	0.87	1.4	5.9	1.9	0.48
	PD	10%	65%	4%			8 %	83%	54%		
486	**In-situ**	0.97	0.18	1.3	1.1	0.02	0.91	0.55	1.5	0.9	0.04
	Satellite	0.96	0.31	1.5	1.2	0.03	0.92	0.53	1.5	1	0.02
	PD	1%	42%	13%			10%	4%	0%		
		486 nm difference relative to 443 nm (regression performance in %)									
486 vs. 443	**In-situ**	+2%	-60%	-41%			-3%	-4%	-44%		
	Satellite	+11%	-75%	-28%			+6%	-62%	-75%		

Parameters:

N: Number of data points
R^2: Coef. Of determination
ε : Error, sum of orthogonal distances.
RMSE: Root mean square error
Int.: Intercept
PD: percentage difference.

Color code shows

Reduced regression performance between in-situ and satellite chlorophyll-*a* retrievals when using R_{rs} 443.

Increased regression performance between in-situ and satellite chlorophyll-*a* retrievals when using R_{rs} 486.

Improved regression performance for both in-situ and satellite chlorophyll-*a* retrievals when using R_{rs} 486 instead of 443.

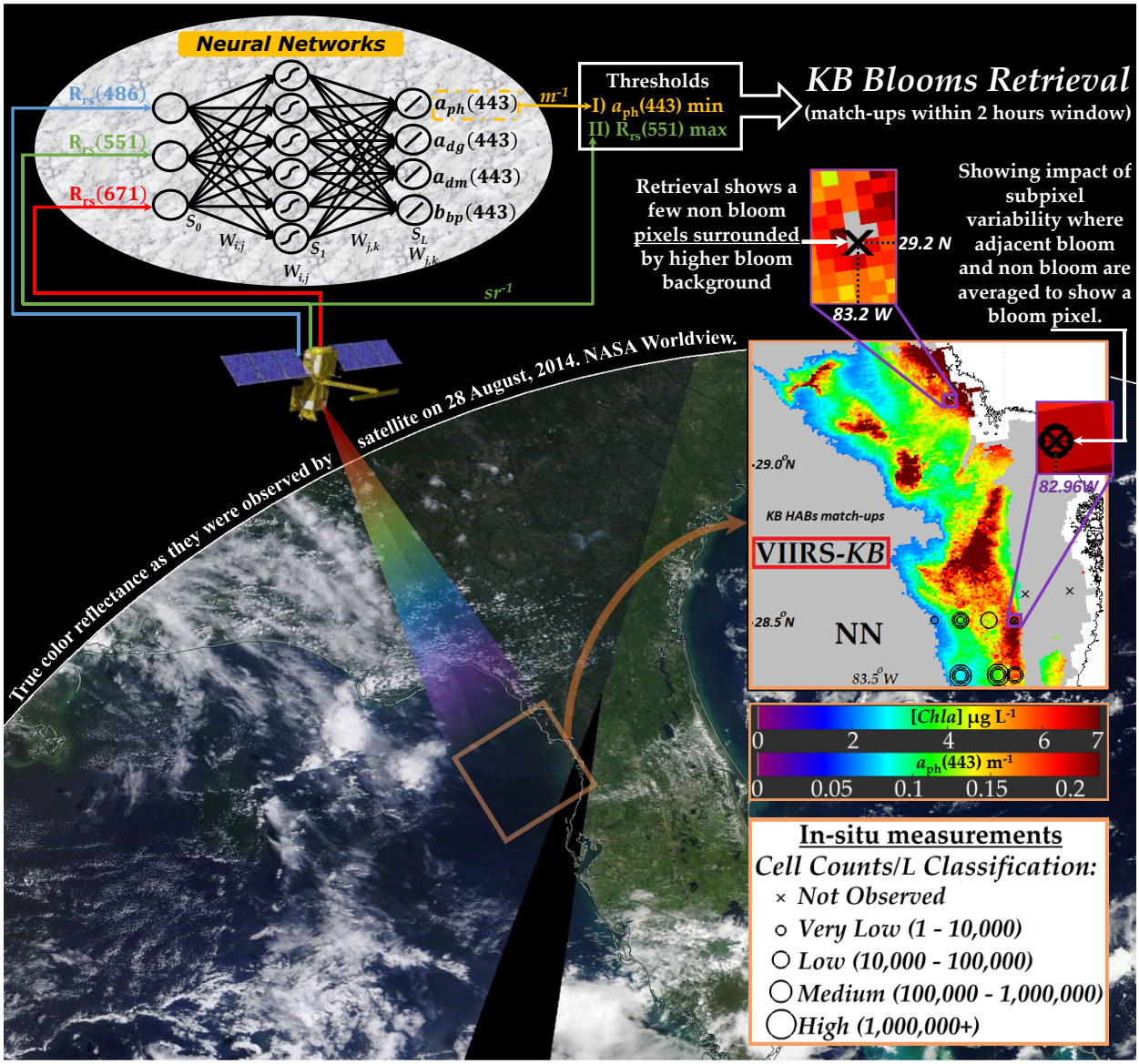

Figure 5 *Example of Karenia brevis harmful algal blooms (KB HABs) retrieval and their corresponding true color imagery from VIIRS satellite in the WFS region using the NN approach.*

In this image, NN algorithms using R_{rs} values from the 486, 551 and 671 nm VIIRS bands are used to retrieve an image of a_{ph443} values in the WFS. Then, additional limiting constraints are applied, in two filter processes, requiring a_{ph443} to be above a limiting value, and backscatter and hence R_{rs} 551 to be below a limiting value, thus eliminating from the retrieved a_{ph443} image all pixels which are not compatible with the existence of *KB* HABs (shown on gray color). The residual image then shows only retrieved a_{ph443} values and their equivalent [*Chla*] values that are consistent with the existence of *KB* HABs. Note that the retrieved [*Chla*] image (shown on the right of this image) is overlaid with *KB* cell counts measurements. The zoomed result on the top highlights the retrieval accuracy where successful retrieval of non-bloom pixel surrounded by high bloom pixels achieved. The impact of sub-pixel variability is also evident on the zoomed top right image where adjacent bloom and non-bloom within the pixel are averaged to show bloom pixels.