Published in final edited form as:
Harmful Algae. 2010 June 1; 9(5): 440–448. doi:10.1016/j.hal.2010.02.002.

Satellite remote sensing of harmful algal blooms: A new multi-algorithm method for detecting the Florida Red Tide (Karenia brevis)

Gustavo A. Carvalho[1,2,*], **Peter J. Minnett**[1,2], **Lora E. Fleming**[2,3], **Viva F. Banzon**[1], and **Warner Baringer**[1]

[1]University of Miami - Rosenstiel School of Marine and Atmospheric Science, Division of Meteorology and Physical Oceanography, 4600 Rickenbacker Causeway, Miami, FL 33149.

[2]NSF NIEHS Oceans and Human Health Center, University of Miami - Rosenstiel School of Marine and Atmospheric Science, 4600 Rickenbacker Causeway, Miami, FL 33149.

[3]University of Miami - Miller School of Medicine, Department of Epidemiology and Public Health, 1120 NW 14th Street, CRB Building (Room 1049), Miami, FL 33136.

Abstract

In a continuing effort to develop suitable methods for the surveillance of Harmful Algal Blooms (HABs) of Karenia brevis using satellite radiometers, a new multi-algorithm method was developed to explore whether improvements in the remote sensing detection of the Florida Red Tide was possible. A Hybrid Scheme was introduced that sequentially applies the optimized versions of two pre-existing satellite-based algorithms: an Empirical Approach (using water-leaving radiance as a function of chlorophyll concentration) and a Bio-optical Technique (using particulate backscatter along with chlorophyll concentration). The long-term evaluation of the new multi-algorithm method was performed using a multi-year MODIS dataset (2002 to 2006; during the boreal Summer-Fall periods – July to December) along the Central West Florida Shelf between 25.75°N and 28.25°N. Algorithm validation was done with in situ measurements of the abundances of K. brevis; cell counts $\geq 1.5\times10^4$ cells l^{-1} defined a detectable HAB. Encouraging statistical results were derived when either or both algorithms correctly flagged known samples. The majority of the valid match-ups were correctly identified (~80% of both HABs and non-blooming conditions) and few false negatives or false positives were produced (~20% of each). Additionally, most of the HAB-positive identifications in the satellite data were indeed HAB samples (positive predictive value: ~70%) and those classified as HAB-negative were almost all non-bloom cases (negative predictive value: ~86%). These results demonstrate an excellent detection capability, on average ~10% more accurate than the individual algorithms used separately. Thus, the new Hybrid Scheme could become a powerful tool for environmental monitoring of K. brevis blooms, with valuable consequences including leading to the more rapid and efficient use of ships to make in situ measurements of HABs.

* Corresponding author: tel.: +1.305.421.4104; fax: +1.305.421.4622 gcarvalho@rsmas.miami.edu .

Keywords

Algorithm development; Central West Florida Shelf (Gulf of Mexico); Detection; Harmful Algal Bloom (HAB); Hybrid Scheme; Florida Red Tide (Karenia brevis); Ocean color (MODIS-Aqua); Satellite remote sensing

1. Introduction

A host of negative impacts can result when algal concentrations exceed background levels particularly when the algal species are capable of producing toxic substances, often know as "harmful algal blooms" (HABs) or as "red tides" (Smayda 1997a;1997b). Due to the apparent worldwide increase of HABs over recent years, the study of the possible effects of certain HABs on human health has spurred public attention from the media and society at large (e.g., Fleming and Laws, 2006;Schrope, 2008).

The West coast of Florida (Figure 1) often experiences severe HABs caused by the toxic dinoflagellate, Karenia brevis, which produces potent neurotoxins called brevetoxins (Steidinger, 2009). Human health can be effected during K. brevis blooms if brevetoxin-contaminated shellfish are consumed and/or the aerosolized toxin is inhaled (Kirkpatrick et at., 2004;Fleming et al., 2007). When present, blooms of K. brevis may also poison sea life, such as seabirds, fish and marine mammals (Landsberg, 2002;Shumway et al., 2003). Serious localized economic losses have also been linked with the "Florida Red Tide" (Hoagland et al., 2009). Because of the recurring and undesirable consequences caused by K. brevis blooms, monitoring of water quality for the occurrence of HABs is a continuing and pressing need in this region, and thereby a fundamental requirement for mitigation efforts and sustainable resource management (Dyble et al., 2008;Heil and Steidinger, 2009).

Improved capabilities for monitoring HABs with high temporal and spatial resolution have long been desired (McCluney, 1974; Murphy et al., 1975). Indeed, there is a widespread interest in the ability to locate and forecast HABs in the world's coastal areas (e.g., Sacau-Cuadrado et al., 2003; Olascoaga et al., 2008). Early warning systems using autonomous underwater vehicles, prototypes of sentinel instruments, real-time monitoring, and other approaches such as moorings, profiling floaters, drifting buoys, and sophisticated numerical modeling have all added high-quality contributions to traditional environmental research and monitoring efforts (e.g., Lee et al., 2005; Robbins et al., 2006). However, most in situ monitoring programs are affected by spatial and temporal biases due to sparse non-uniform coverage. Thus, many questions regarding the origin and development mechanisms of the Florida Red Tide (and other HABs elsewhere) remain unanswered (Tester et al., 1993, Walsh and Steidinger, 2001, Brand and Compton, 2007).

Since their inception, satellites carrying ocean color instruments proved to be a valuable complementary tool for synoptically monitoring trends of oceanographic features associated with K. brevis blooms (e.g., Haddad, 1982). Throughout the last decade, an analyst-dependent HAB Forecasting System (http://tidesandcurrents.noaa.gov/hab) has been used by the National Oceanic and Atmospheric Administration (NOAA) to detect potential blooms of K. brevis in the Gulf of Mexico using data of the Sea-viewing Wide Field-of-view Sensor (SeaWiFS; McClain et al., 2004). In short, the conceptual basis of this "Operational Method" lies in using satellite-measured chlorophyll (Chl) anomalies computed relative to the average of the previous two months (with a window of two weeks to the start of the mean calculation); an anomaly of 1 mg m^{-3} indicates a likely increase of 10×10^4 cells l^{-1} in the concentration of K. brevis (Stumpf et al., 2003; Tomlinson et al., 2004; 2009).

More recently, two other satellite-based algorithms have also shown promise at the remote detection of K. brevis blooms. Both are applied to single images, therefore avoiding the use to construct a reference field. The Empirical Approach, introduced by Carvalho (2008) and Carvalho et al. (2010), identifies HABs when the magnitude of the satellite-measured water leaving radiance ($L_{wSAT}(\lambda{\sim}550)$) is lower than the value of modeled total scatter ($b_{MOREL}(550)$; Morel, 1988), as a function of satellite-measured Chl. The Bio-optical Technique, developed by Cannizzaro (2004) and Cannizzaro et al. (2008), employs a similar approach but uses two independent criteria that should occur simultaneously for indicating a pixel as HAB-positive: the satellite-measured Chl value must be high ($\geq$1.5 mg m^{-3}) at the same time as the value of a more highly-derived satellite product, i.e., the particulate backscatter ($b_{bpSAT}(\lambda{\sim}550)$), should be less than the calculated value for this product using Morel (1988) model ($b_{bpMOREL}(550)$). In both algorithms, pixels not satisfying these conditions were classified as HAB-negative cases. A similar cell count threshold was used to define a detectable HAB in the identifications of the Empirical and Bio-optical algorithms: $\geq 1.5\times10^4$ and $\geq 1.0\times10^4$ cells l^{-1}, respectively; samples containing fewer cells than these thresholds were classified as non-bloom cases.

These two algorithms were based on the same reference wavelength, i.e.,$\lambda{\sim}$550nm. While the Bio-optical Technique was developed with a shipboard optical dataset but tested originally with one SeaWiFS image, which has λ=555nm as its closed band (Cannizzaro et al, 2008), the Empirical Approach was developed using data from MODIS (MODerate Resolution Imaging Spectroradiometer; Esaias et al., 1998) where the closest band is centered at λ=551nm. At these wavelengths, total absorption is small and $b_{bp}(\lambda{\sim}550)$ is the major attenuation facto allowing factor allowing one to estimate $b_{bp}(\lambda{\sim}550)$ from $L_w(\lambda{\sim}550)$ measurements (Pope and Fry, 1997; IOCCG, 2006).

In a preliminary investigation, Carvalho (2008) demonstrated that the Empirical Approach was able to produce spatially-coherent outcomes comparable to the HAB detection produced in the original Bio-optical Technique application of Cannizzaro et al. (2008) – a field-verified bloom of K. brevis at the end of August 2001 was successfully identified in a multi-sensor case of study using both algorithms applied to data of both SeaWiFS and MODIS. The application of the Bio-optical Technique to a long-time series of SeaWiFS data did not produce persuasive results (Hu et al., 2008), but this technique has subsequently been successfully applied to MODIS data in conjunction with data from three ~5-day cruises in 2005 and 2006 (Cannizzaro et al., 2009). In the study of Carvalho et al. (2010), an analysis of the Bio-optical Technique was carried out using a nearly five-year MODIS dataset from 2002 to 2006; results suggested that this technique could perform well provided the satellite dataset was carefully quality controlled and considering data distribution with seasonal and latitudinal restrictions.

Carvalho et al. (2010) also completed a long-term comparative investigation of the performances of the Empirical Approach, the Bio-optical Technique and the Operational Method. With an automated pixel-based match-up analysis that used the same in situ dataset for validation of the three algorithms, the Operational Method did not identify distinct populations of HABs and non-blooms as effectively as did the other two algorithms using MODIS data. When cell counts were plotted against the optical parameters used in the algorithms, the two populations produced by the Operational Method overlapped significantly (refer to Figures 7.10 through 7.18 in Carvalho (2008)). The Empirical and Bio-optical algorithms could be optimized by varying the appropriate threshold values – this was done with geographical and seasonal restrictions – and this resulted in modified versions which had improved performance in differentiating HABs from non-blooming conditions – summarized in Table 1 [panels X-XII]. The user was left with a choice of using one or the other of the optimized algorithms.

There is still scope for improving HAB detection with satellite measurements, and the objective of the current study was to determine whether the sequential application of the optimized versions of the Empirical Approach and the Bio-optical Technique would lead to improvements in identifying the optical characteristics related with the Florida Red Tide. The combination of these two algorithms has not been attempted previously, and represents a new approach to the problem. Thus, by exploiting the detection capabilities of the Empirical and the Bio-optical algorithms, a new method to detect K. brevis blooms using satellite measurements was explored.

2. Methods

In the current study, the methodology used to establish the long-term detection capabilities of the optimized versions of the Empirical Approach and the Bio-optical Technique was followed (Carvalho et al., 2010). This involved performing an analysis with the same rigorous quality control to produce the "valid cloud-free match-ups" between satellite and in situ measurements (Figure 2). Shipboard and satellite measurements were matched in space (i.e., the pixel containing the in situ measurement) and time (i.e., in situ data within the same day of the satellite overpass). The rationale behind using only satellite measurements from the pixel containing the position at which the sample was taken avoids uncertainties from the use of larger areas including neighboring pixels to the field location, while the use of this time-acceptance window was justified by the slow growth rate of K. brevis cells (i.e., between 0.05 to 0.5 doublings day^{-1}; Shanley and Vargo, 1993). As the optimal accuracies of the Empirical and Biooptical algorithms were temporally and spatially constrained (Carvalho et al., 2010), the data selected for the current investigation were taken only during the boreal Summer-Fall periods (i.e., July to December) along the region of the Central West Florida Shelf (CWFS; south of 28.25°N close to Tampa Bay and north of 25.75°N near Cape Romano; Figure 1).

A comprehensive in situ database containing historical and current K. brevis monitoring data (Haverkamp et al., 2004) was used to validate the outcome of the new multi-algorithm method; the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute provided this database. As this database has been previously used for validating the long-term performance of the Empirical, Bio-optical and Operational algorithms (Carvalho et al., 2010), the same cell count threshold was used to evaluate the accuracy of the new multialgorithm. A binary classification sorted the database for the presence of Florida Red Tide (i.e., "red tide"; samples containing $\geq 1.5\times10^4$ cells of K. brevis l^{-1}) and for the absence of HABs (i.e., "non-blooming" conditions; samples with $<1.5\times10^4$ cells of K. brevis l^{-1}).

The satellite imagery used was obtained from the Ocean Color Website (http://oceancolor.gsfc.nasa.gov/cgi/browse.pl?sen=am) for the period of July 2002 to mid October 2006. All level-1A MODIS-Aqua scenes (L1A files), covering the region of interest (Figure 1), were processed to level 2 (L2 files) with SeaDAS version 5.0.3 using default settings of the MSL12,4 version 5.2.3 (http://oceancolor.gsfc.nasa.gov/seadas). A large library of geophysical parameters derived from the satellite data was generated: total scatter($b(\lambda)$); particulate backscatter ($b_{bp}(\lambda)$); total absorption ($a(\lambda)$); absorption due to detritus and colored organic dissolved matter together ($a_{dg}(\lambda)$); and absorption due to phytoplankton ($a_{ph}(\lambda)$); these inherent optical properties (IOPs) of sea-water were derived with the Carder semi-analytical algorithm (Carder et al., 1999). This library also included: water-leaving radiance ($L_w(\lambda)$); normalized water-leaving radiance ($nL_w(\lambda)$); remote-sensing reflectance ($R_{rs}(\lambda)$); and were produced in 9 bands centered at $\lambda\sim$ 412, 443, 488, 531, 551, 667, 678, 748 and 869nm. In addition, Chl (derived with the global OC3M band-ratio algorithm; O'Reilly et al., 2000); downwelling diffuse attenuation coefficient (K_{490}); fluorescence line-height (FLH); sea-surface temperature (SST); and satellite zenith angle (θ) were included. Figure 2 shows the flowchart illustrating the processes to obtain the pairs of valid match-ups to determine the

accuracy, and test the performances, of the specific HAB-detection algorithms applied to satellite imagery. More detail is given elsewhere (Carvalho, 2008;Carvalho et al., 2010).

3. Statistical assessment

The ability to differentiate between HABs and non-blooming conditions, i.e., the satellite-detection effectiveness, was evaluated using basic statistical metrics (Congalton, 1991; Sokal and Rohlf, 1995; Mangnusson and Mourão, 2004), as described below. Following Carvalho et al. (2010), these measures were also used in the current study to evaluate the accuracy of the new multi-algorithm method; and in a comparison with the accuracies of the individual optimized versions of the Empirical Approach and the Bio-optical Technique. Error matrices (Table 1 [panel I-III]) give details of some applicable quantities used in the process of evaluating the accuracies of the algorithms derived from the numbers of: a priori known red tide samples correctly flagged as HAB-positive (designated as A); a priori known red tide samples incorrectly flagged as HAB-negative (B); a priori known non-bloom samples incorrectly flagged as HAB-positive (C); and a priori known non-bloom samples correctly flagged as HAB-negative (D) – see Table 2.

A standard and widely-used metric to assess the accuracy of classification algorithms is the overall accuracy, i.e., total of successful identifications divided by the total population ([A+D]/[A+B+C+D]; Table 2). However, the user can be misled if this metric is used alone; thus other basic metrics should be considered collectively. To help to understand these metrics they were divided in two groups; as shown in Table 1 [panels II and III].

The first set of metrics evaluated how many in situ samples, among the a priori known observed presence or absence of K. brevis blooms, were correctly identified by the algorithms, so as it gave the rate of failure in the detection (Table 1 [panel II]). Thus, four questions should be answered:

Question 1a. What portion of the red tide population was correctly flagged as HAB-positive (i.e., A/[A+B])? This is the sensitivity of the algorithm, also called as producer's accuracy (Congalton, 1991).

Question 1b. Among the red tide population, how many samples were not correctly detected by the algorithm (i.e., B/[A+B])? This is the fraction of false negatives.

Question 1c. Among the non-bloom population, how many samples were wrongly classified by the algorithm (i.e., C/[C+D])? This is the fraction of false positives.

Question 1d. What portion of the non-bloom population was correctly identified as HAB-negative (i.e., D/[C+D])? This is the specificity of the algorithm.

The second set of metrics evaluated the accuracies of the algorithms (i.e., the predictive values; Table 1 [panel III]). These metrics are important as they represent the probability of the algorithm correctly identifying observed in situ conditions (i.e., the actual satellite identification based on the presence or absence of K. brevis blooms determined by microscopic analysis of the water samples). Hence, the answer to four other questions were sought:

Question 2a. What portion flagged as HAB-positive was represented by red tide samples (i.e., A/[A+C])? This is the positive predictive value, also known as user's accuracy (Congalton, 1991).

Question 2b. Among the cases flagged as HAB-positive, how many samples were in fact non-blooms (i.e., C/[A+C])? This is the inverse of the positive predictive value.

Question 2c. Among the cases identified as HAB-negative, how many samples were in fact red tides (i.e., B/[B+D])? This is the inverse of the negative predictive value.

Question 2d. What portion identified as HAB-negative was represented by non-bloom samples (i.e., D/[B+D])? This is the negative predictive value.

Predictive values are strongly dependent on the prevalence of red tides (i.e., the number of known red tide samples over the total population; [A+B]/[A+B+C+D]; Table 2). One should note that ideal algorithms should have both of the inverse, positive and negative of the predictive values equal to zero. Also, an ideal algorithm should have no cases of false negatives nor false positives.

4. The Hybrid Scheme

Despite the satisfactory performances of the optimized version of the Empirical Approach and the Bio-optical Technique (Table 1 [panels X-XII]), there were still some samples that were correctly identified by one algorithm, but not by the other; and vice-versa (Carvalho et al., 2010). Therefore, we hypothesized that HAB-detection improvements could be possible with the sequential application of the Empirical and Bio-optical algorithms into a single multi-algorithm method. Thus, we proposed a Hybrid Scheme dedicated to detect K. brevis blooms in the CWFS Summer-Fall using satellite ocean color imagery.

The new Hybrid Scheme was applied in two configurations; as depicted by the color-code in Figure 3a. In the 1st configuration of the Hybrid Scheme, a positive identification occurred when either or both algorithms correctly flagged known samples (Figure 3a: YELLOW +BLUE), while both algorithms had to fail simultaneously in detecting a known sample to establish false conditions (Figure 3a: WHITE). To determine the detection accuracy of the 2nd configuration of the Hybrid Scheme, both algorithms had to flag simultaneously known samples correctly for a positive identification (Figure 3a: BLUE), while false conditions occurred when either or both algorithms failed to classify a known sample correctly (Figure 3a: YELLOW+WHITE).

The two configurations of the Hybrid Scheme were compared to quantify whether improved detection of the K. brevis blooms could be achieved compared to the separate use of the Empirical Approach and the Bio-optical Technique, and if so, proposed as a potential remote-sensing monitoring tool. The basic statistical assessments (as described by the eight questions of Section 3) were calculated individually at the two configurations of the Hybrid Scheme.

5. Results

Of the 371 in situ samples of the CWFS Summer-Fall that passed the match-up analysis quality control tests (Figure 2), 145 counted as red tides while the rest were non-bloom cases (n=226). The detection frequency observed by the identification of red tide and non-blooming conditions in the satellite data are shown in Figure 3 [b, c] for the individual optimized versions of the Empirical and Bio-optical algorithms, and in Table 2 for both configurations of the Hybrid Scheme. The answers to the first set of four questions (Section 3: 1a-1d) can be found in Table 1 [panels V, VIII, and XI], while the answers to the second set of four questions (Section 3: 2a-2d) can be found in Table 1 [panels VI, IX, and XII].

In terms of the observed incidence of red tide, established by the cell count threshold of 1.5×10^4 cells l^{-1}, both configurations of the Hybrid Scheme had equal values: prevalence of red tides of ~39%; n=145/371 (Table 2). On the other hand, the results of the overall accuracy varied (Table 2), as this metric was determined by the total number of correct samples each configuration detected: the 1st configuration of the Hybrid Scheme had a good value (~79%; n=293/371), while the 2nd configuration had a moderate value (~62%; n=227/371). In comparison, the overall accuracies of the individual optimized versions of the Empirical and

Bio-optical algorithms were between these values (~70%; refer to Table 6 in Carvalho et al. (2010)).

The 1st configuration of the Hybrid Scheme had better detection rates than the 2nd configuration (Table 1). The sensitivity (80%; n=116/145) and specificity (78%; 177/226) observed for the 1st configuration were good, and were accompanied by low levels of false negatives (20%; n=29/145) and false positives (22%; n=49/226); as shown in Table 1 [panel V]. Poorer sensitivity (62%; 90/145) and specificity (61%; n=137/226) were observed at the 2nd configuration along with more cases of false negatives (38%; n=55/145) and false positives (39%; n=89/226); as shown in Table 1 [panel VIII]. The long-term detection accuracies of the individual optimized versions of the Empirical Approach and Bio-optical Technique, derived in Carvalho et al. (2010), are summarized in Table 1 [panel XI].

As can be seen in Table 2, of the total population (n=371), the 1st configuration of the Hybrid Scheme flagged 165 as HAB-positive ([A=116] plus [C=49]), while 206 samples were classified as HAB-negative ([B=29] plus [D=177]); Table 1 [panel IV]. In the case of the 2nd configuration, 179 samples were flagged as HAB-positive ([A=90] plus [C=89]), while 192 samples were classified as HAB-negative ([B=55] plus [D=137]); Table 1 [panel VII]. Thus, the derived positive predictive value of the 1st configuration of the Hybrid Scheme (Table 1 [panel VI]) was satisfactorily good (70%; n=116/165), so was the negative predictive value (86%; n=177/206). The excellent predictive values of the 1st configuration did not translate into the corresponding values of the 2nd configuration, which was less accurate (Table 1 [panel IX]). The predictive values of the individual optimized versions of the Empirical and Bio-optical algorithms (Carvalho et al., 2010) are summarized in Table 1 [panel XII].

6. Discussion

The identification of HABs using satellite measurements and the establishment of the reliability of such retrievals are challenging tasks, largely because of the limited information in the satellite data, the problems of validating satellite ocean color retrievals in near-shore areas and the associated difficulties in finding a representative field dataset (Bailey and Wang, 2001; Werdell et al., 2003). Nevertheless, in this study, the use of the in situ database maintained by the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute successfully achieved the purpose of validating the detection of K. brevis blooms containing $\geq 1.5\times10^4$ cells l^{-1} using MODIS measurements.

Although the relatively large native surface resolution of MODIS ocean color bands (of 1 km^2 at the nadir point) along with the fairly relaxed time-acceptance match-up window (in situ and satellite data within the same day) may have added some uncertainties to the analysis, the stringent exclusion criteria, which retained only pairs as free as possible of undesired artifacts, greatly improved the validity of our results (Figure 2). The final cloud-free match-ups (n=[371/3420] ~11%; comprising ~95 samples $year^{-1}$) served as a dataset for determining the long-term accuracy of the new Hybrid Scheme in the CWFS Summer-Fall. The number of red tide samples and the total population used in the current study ([n=145 and n=371; respectively]; Tables 1 and 2) was larger than those used in other related published studies to detect K. brevis blooms along the CWFS using satellite measurements – e.g., [n=42 and n=119; respectively] refer to Table 2 in Tomlinson et al. (2004); and [n=27 and n=90; respectively] refer to Table 3 in Cannizzaro et al. (2009).

The use of 2-by-2 tables to display the basic statistical results has been done to help the interpretation of the metrics used to establish the accuracy of the satellite-based algorithms (Table 1). Requiring the Empirical Approach and the Bio-optical Technique to flag correctly a sample at the same time (Figure 3a), the 2nd configuration of the Hybrid Scheme produced

modest detection results (i.e., sensitivity and specificity) with almost twice the number of false negatives and false positives than those of the 1st configuration (Table 1 [panels V and VIII]). This pattern was also observed in their predictive values (Table 1 [panels VI and IX]). Thus, after answering the eight questions of Section 3, the comparison of the HAB detection accuracies of both configurations of the Hybrid Scheme with the accuracies of the individual optimized version of the Empirical and the Bio-optical algorithms (Table 1 [panels X-XII]) revealed that the 1st configuration had better detection qualities (being in average ~10% superior; Table 1 [panels IV-VI]) while the 2nd configuration was not as accurate, being roughly ~10% worse (Table 1 [panels VII-IX]) than the individual algorithms used separately.

The strength of the Hybrid Scheme lies in the use of a multi-parameter approach (i.e., $L_w(\lambda{\sim}550)$, $b_{bp}(\lambda{\sim}550)$ and Chl), as the identification of K. brevis blooms was not solely derived from Chl retrievals. Therefore, the effectiveness of the 1st configuration of the Hybrid Scheme in correctly flagging red tide samples as HABs-positive was observed together with the ability to classify correctly as HAB-negative most of the non-blooming conditions, whether at low-Chl or at high-Chl concentrations dominated by other non-toxic algal species. Thus, the good detection accuracy of the 1st configuration suggests that the Hybrid Scheme could be an asset for reliable environmental monitoring. Indeed, it can be used to direct field-monitoring campaigns improving the timeliness of in situ sampling of HABs. This would assist scientists to investigate the Florida Red Tide in a more focused manner, thus generating better understanding of major issues related with the occurrence of HABs, e.g., the mechanisms that trigger, perpetuate and disperse blooms (Kusek et al., 1999; Smayda, 2007).

It is important to bear in mind that sensitivity and specificity are independent variables, as are false negatives and false positives; however, sensitivity is coupled with false negatives, and cases of false positives are the complements of specificity. From a resource manager's standpoint, an error in flagging a HAB (i.e., a false negative case) is not desirable given possible human and environmental health impacts (e.g., a failure in public protection may occur if shellfish beds are contaminated). The failure to identify correctly non-blooming conditions (i.e., false positive cases) is not financially acceptable since no HABs are present and erroneous announcements may have large localized economic consequences (e.g., loss of tourism). It can be argued under variations of the "Precautionary Principle" (Kriebel et al., 2001) that in the case of false positive errors, at least undesirable public health risks are avoided, which would be a more socially acceptable alternative to be integrated into decision-making management and strategies.

At the present time, the capabilities of the ocean color satellite sensors are not yet able to provide sufficient information for a confident identification of HABs species. Neither can the algorithms determine a decision for the closure of recreational and commercial shellfish beds based on remote sensing data alone as the reference threshold of $1.5{\times}10^4$ cell l^{-1} is still above the guidelines used in monitoring and regulatory decisions (i.e., $0.5{\times}10^4$ cell l^{-1}; FDA, 2005). However, the reference level used in the current study to define detectable K. brevis bloom within the satellite measurements is more realistic and more sensitive than other values found in the published literature (e.g., $10{\times}10^4$ cells $l^{\times 1}$ in Tester et al. (1998) – this is also the same threshold implemented in the Operational Method; Fisher et al., 2005).

On the other hand, as good monitoring capabilities were demonstrated in the current analysis, the uncertainties in detecting HABs with satellite imagery could indeed be reduced by using the Hybrid Scheme in the 1st configuration (where either or both algorithms correctly detecting a sample positively). Thus, resource managers or scientists studying HABs would have a better chance to be aware of, or to locate sites suspected of having the presence or absence of blooms of K. brevis based on the outcomes of the 1st configuration of the Hybrid Scheme. Similarly, beach-goers would have a more confidence in avoiding potentially contaminated areas. And,

as the new multi-algorithm method achieved potential for scientific and operational applications, the Hybrid Scheme could assist in accurately directing field monitoring, which would promote improved seafood safety action and reducing the risk of Neurotoxic Shellfish Poisoning (Watkins et al., 2008).

Even though K. brevis is believed to be mainly restricted to Gulf of Mexico waters, the genus Karenia can be found in different oceans and seas around the globe (Haywood et al., 2004; Steidinger et al., 2008). Therefore, the detection capabilities presented by the Hybrid Scheme could also be applicable, in principle, to other regions where blooms of Karenia species may impart the same optical characteristics to the water (Schofield et al., 1999).

7. Concluding remarks

The sequential application of the optimized versions of the Bio-optical Technique and the Empirical Approach (Carvalho et al., 2010) produced a new multi-algorithm method specifically dedicated to detect the Florida Red Tide using remote sensing measurements – this was the first-time application that this new method, referred to as Hybrid Scheme, has been reported. The validity of the Hybrid Scheme was verified with a comprehensive in situ database of K. brevis cell abundances along the Central West Florida Shelf during the Summer-Fall periods (i.e., July to December) using a multi-year MODIS-Aqua dataset (2002 to 2006).

With an automated analysis, the long-term performance of the Hybrid Scheme was evaluated in two configurations. Better detection accuracies for K. brevis blooms containing $\geq 1.5\times10^4$ cells 1^{-} were observed in the 1st configuration (i.e., equally good sensitivity and specificity (~80%) with simultaneously good predictive values (>70%); Tables 1 and 2; Figure 3). The 1st configuration of the Hybrid Scheme benefited from the strengths of the individual algorithms – when the Empirical Approach missed a red tide there was still a chance of the same sample being correctly flagged by the Bio-optical Technique, and vice-versa.

The application of the 1st configuration of the Hybrid Scheme could provide a valuable low-cost means of HAB surveillance. Resource managers would be able to inform the general public about the location of HABs as they occur, adding confidence to mitigation action and reducing local economic impacts. Moreover, the accurate detection of K. brevis blooms could assist researches in the early identification of HABs as well as helping to create more appropriate public guidance and medical health plans.

Acknowledgments

We would like acknowledge the contributions to this study by Edward Kearns, Robert Evans, Jennifer Cannizzaro, Kendall Carder, Chuanmin Hu, Sharon Smith, Julie Hollenbeck, Maria Villanueva, Gan Changlin and Christina Plattner. The historical in situ database maintained by the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute, and provided by Cynthia A. Heil, made this work possible. Financial support was provided by NSF and NIEHS Oceans and Human Health Center grants: NSF#OCE0432368/0911373 and NIEHS#P50-ES12736.

References

Bailey, S.; Wang, M. In situ aerosol optical thickness collected by the SIMBIOS program (1997-2000): Protocols, and data QC and analysis. NASA Goddard Space Flight Center; Greenbelt, MD: 2001. Satellite aerosol optical thickness match-up procedures; p. 70-72.NASA/TM-2001-209982

Bailey SW, Werdell PJ. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. Remote Sens. Environ 2006;102:12–23.

Brand LE, Compton A. Long-term increase in Karenia brevis abundance along the Southwest Florida Coast. Harmful Algae 2007;6:232–252. [PubMed: 18437245]

Cannizzaro, JP. M.S. Thesis. University of South Florida at St. Petersburg; FL: Mar 29. 2004 Detection and quantification of Karenia brevis blooms on the West Florida Shelf from remotely sensed ocean color imagery; p. 71

Cannizzaro JP, Carder KL, Chen FR, Heil CA, Vargo GA. A novel technique for detection of the toxic dinoflagellate Karenia brevis in the Gulf of Mexico from remotely sensed ocean color data. Cont. Shelf Res 2008;28(1):137–158.

Cannizzaro JP, Hu C, English DC, Carder KL, Heil CA, Muller-Karger FE. Detection of Karenia brevis blooms on the west Florida shelf using in situ backscattering and fluorescence data. Harmful Algae 2009;8:898–909.

Carder KL, Chen FR, Lee ZP, Hawes SK, Kamykowski D. Semi-analytic Moderate-Resolution Imaging Spectrometer algorithms for chlorophyll a and absorption with bio-optical domains based on nitrate-depletion temperatures. J. Geophys. Res 1999;104:5403–5421.

Carvalho, GA. M.S. Thesis. University of Miami RSMAS-MPO at Miami; FL: Apr 11. 2008 The use of satellite-based ocean color measurements for detecting the Florida Red Tide (Karenia brevis); p. 174Available at http://etd.library.miami.edu/theses/available/etd-04162008-210856

Carvalho GA, Minnett PJ, Banzon VF, Baringer W, Heil CA. Long-term evaluation of three satellite ocean color algorithms for identifying harmful algal blooms (Karenia brevis) along the west coast of Florida. Remote Sens. Environ. 2010 (submitted).

Congalton RG. A review of assessing the accuracy of classification of remote sensed data. Remote Sens. Environ 1991;37:35–46.

Dyble J, Bienfang P, Dusek E, Hitchcock G, Holland F, Laws E, Lerczak J, McGillicuddy DJ Jr. Minnett PJ, Moore SK, O'Kelly C, Solo-Gabriele H, Wang JD. Environmental controls, oceanography and population dynamics of pathogens and harmful algal blooms: Connecting sources to human exposure. Environ. Health 2008;7(2):S5. doi:10.1186/1476-069X-7-S2-S5. [PubMed: 19025676]

Esaias W, Abbott M, Barton I, Brown OB, Campbell JW, Carder KL, Clark DK, Evans RH, Hoge FE, Gordon HR, Balch WM, Letelier R, Minnett PJ. An overview of MODIS capabilities for ocean science observations. IEEE T. Geosci. Remote 1998;36(4):1250–1265.

FDA (Florida Department of Agriculture and Consumer Services). National Shellfish Sanitation Program: Guide for the Control of Molluscan Shellfish. 2005.

Fisher, KM.; Allan, AL.; Keller, HM.; Bronder, ZE.; Fenstermacher, LE.; Vincent, MS. Annual repost of the Gulf of Mexico harmful algal bloom operational forecast system (GOM HAB-OFS). NOAA Technical Repost NOS CO-OPS 047. 2005.

Fleming LE, Laws E. Overview of the ocean and human health special issue. Oceanogr 2006;16(2):18–23.

Fleming LE, Kirkpatrick B, Backer LC, Bean JA, Wanner A, Reich A, Zaias J, Cheng YS, Pierce R, Naar J, Abraham WM, Baden DG. Aerosolized red-tide toxins (Brevetoxins) and asthma. CHEST 2007;131(1):187–194. [PubMed: 17218574]

Gordon HR, McCluney WR. Estimation of the depth of sunlight penetration in the sea for remote sensing. Applied Optics 1975;14(2):413–416. [PubMed: 20134900]

Hoagland P, Jin D, Polansky LY, Kirkpatrick B, Kirkpatrick G, Fleming LE, Reich A, Watkins SM, Ullman SG, Backer LC. The Costs of Respiratory Illnesses Arising from Florida Gulf Coast Karenia brevis Blooms. Environ. Health Persp 2009;117:1239–1243.

Haddad, KD. M.S. Thesis. University of South Florida at St. Petersburg; FL: Apr. 1982 Hydrographic factors associated with west Florida toxic red tide blooms: an assessment for satellite prediction and monitoring; p. 155

Haverkamp, D.; Steidinger, KA.; Heil, CA. HAB Monitoring and Databases: The Karenia brevis example. In: Hall, S.; Etherrdge, S.; Anderson, D.; Kleindinst, J.; Zhu, M.; Zou, Y., editors. Harmful Algae Management and Mitigation, Asia-Pacific Economic Cooperation (Singapore). 2004. APEC Publication#204-MR-04.2

Haywood AJ, Steidinger KA, Truby EW, Bergquist PR, Bergquist PL, Adamson J, MacKenzie L. Comparative morphology and molecular phylogenetic Analysis of three species of the genus Karenia (Dinophyceae) from New Zeland. J. Phycol 2004;40:165–179.

Heil CA, Steidinger K. Monitoring, management, and mitigation of Karenia blooms in the eastern Gulf of Mexico. Harmful Algae 2009;8:611–617.

Hu C, Luerssen R, Müller-Karger FE, Carder KL, Heil CA. In search of red tides: Observations on the west Florida shelf. Cont. Shelf Res 2008;28(1):159–176.

IOCCG. Remote sensing of inherent optical properties: fundamentals, tests of algorithms, and applications. In: Lee, Z-P., editor. Reports of the International Ocean-Colour Coordinating Group, No. 5. IOCCG; Dartmouth, Canada: 2006. http://www.ioccg.org/reports_ioccg.html/

Kirkpatrick B, Fleming LE, Squicciarini D, Backer LC, Clark R, Abraham W, Benson J, Cheng YS, Johnson D, Pierce R, Zaias J, Bossart GD, Baden DG. Literature review of Florida red tide: implications for human health effects. Harmful Algae 2004;3:99–115. [PubMed: 20411030]

Kriebel D, Tickner J, Epstein P, Lemons J, Levins R, Loechler EL, Quinn M, Rudel R, Schettler T, Stoto M. The Precautionary Principle in Environmental Science. Environ. Health Persp 2001;109(9):871–876.

Kusek KM, Vargo G, Steidinger K. Gymnodinium breve in the field in the lab and in the newspaper - a scientific and journalistic analysis of Florida red tides. Contrib. Mar. Sci 1999;34:1–229.

Landsberg JH. The effects of harmful algal blooms on marine and aquatic organisms. Rev. Fish. Sci 2002;10(2):113–390.

Lee ZP, Carder KL, Arnone RA. Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. Appl. Optics 2002;41(27):5755–5772.

Lee JHW, Hodgkiss IJ, Wong KTM, Lam IHY. Real time observations of coastal algal blooms by an early warning system. Estuar. Coast. Shelf Sci 2005;65:72–190.

Mangnusson, WE.; Mourão, G. Statistics without math. Planta Ed. 2004. p. 136

McClain CR, Feldman GC, Hooker SB. An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series. Deep-Sea Res. Pt II 2004;51:5–42.

McCluney, WR. NASA Red Tide remote sensing research. In: Joyce, E., editor. Proceedings of the Florida Red Tide Conference; 1974. p. 10-11.Fla. Mar. Res. Pub. (8)

Morel A. Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). J. Geophys. Res 1988;93:10,749–10,768.

Murphy EB, Steidinger KA, Roberts BS, Williams J, Jolley JW Jr. An Explanation for the East coast Gimnodinium breve red tide of Noverber 1972. Limnol. Oceanogr 1975;20(3):481–486.

Olascoaga MJ, Beron-Vera FJ, Brand LE, Kocak H. Tracing the early development of harmful algal blooms on the West Florida Shelf with the aid of Lagrangian coherent structures. J. Geophys. Res 2008;113(C12014) doi:10.1029/2007JC004533.

O'Reilly, JE.; Maritorena, S.; Siegel, DA.; O'Brien, MC.; Toole, D.; Mitchell, BG.; Kahru, M.; Chavez, FP.; Strutton, P.; Cota, GF.; Hooker, SB.; McClain, CR.; Carder, KL.; Muller-Karger, F.; Harding, L.; Magnuson, A.; Phinney, D.; Moore, GF.; Aiken, J.; Arrigo, KR.; Letelier, R.; Culver, M. Ocean color chlorophyll a algorithms for SeaWiFS, OC2, and OC4: Version 4. In: Hooker, SB.; Firestone, ER., editors. SeaWiFS Postlaunch Calibration and Validation Analyses. Vol. Vol. 11. NASA Goddard Space Flight Center; Greenbelt, Maryland: 2000. p. 9-23.Part 3. NASA Technical Memorandum 2000-206892

Pope RM, Fry ES. Absorption spectrum (380-700 nm) of pure water. II. Integrating cavity measurements. Appl. Optics 1997;36(33):8710–8723.

Robbins IC, Kirkpatrick GJ, Blackwell SM, Hillier J, Knight CA, Moline MA. Improved monitoring of HABs using autonomous underwater vehicles (AUV). Harmful Algae 2006;5:749–761.

Sacau-Cuadrado M, Conde-Pardo P, Otero-Tranchero P. Forecast of red tides off the Galician coast. Acta Astronaut 2003;53:439–443.

Schrope M. Red tide rising. Nature 2008;452:24–26. [PubMed: 18322500]

Schofield O, Grzymski J, Bissett WP, Kirkpatrick GJ, Millie DF, Moline M, Roesler CS. Optical monitoring and forecasting systems for harmful algal blooms: possibility or pipe dream? J. Psychol 1999;35:1477–1496.

Shanley, E.; Vargo, GA. Cellular composition, growth, photosynthesis, and respiration rates of Gymnodinium breve under varying light levels. In: Smayda, TJ.; Shimizu, Y., editors. Toxic Phytoplankton Blooms in the Sea. Elsevier; New York: 1993. p. 919-923.

Shumway SE, Aslen SM, Boersma PD. Marine birds and harmful algal blooms: sporadic victims or under-reported events? Harmful Algae 2003;2:1–17.

Smayda TJ. Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea. Limnol. Oceanogr 1997a;42(5 part 2):1137–1153.

Smayda TJ. What is a bloom? A commentary. Limnol. Oceanogr 1997b;42(5, part 2):1132–1136.

Smayda TJ. Reflections on the ballast water dispersal - harmful algal bloom paradigm. Harmful Algae 2007;6:601–622.

Sokal, RR.; Rohlf, FJ. Biometry: The Principles and Practice of Statistics in Biological Research. 3rd ed. Freeman; New York: 1995. p. 850

Steidinger KA. Historical perspective on Karenia brevis red tide research in the Gulf of Mexico. Harmful Algae 2009;8(4):549–561. doi:10.1016/j.hal.2008.11.009.

Steidinger KA, Wolny JL, Haywood AJ. Identification of Kareniaceae (Dinophyceae) in the Gulf of Mexico. Nova Hedwigia 2008;133:269–284.

Stumpf RP, Culver ME, Tester PA, Tomlinson MC, Kirkpatrick GJ, Pederson BA, Truby E, Ransibrahmanakul V, Soracco M. Monitoring Karenia brevis blooms in the Gulf of Mexico using satellite ocean color imagery and other data. Harmful Algae 2003;2:147–160.

Tester, PA.; Geesey, ME.; Vukovich, FM. Gymnodinium breve and global warming; what are the possibilities?. In: Smayda, TJ.; Shimizu, Y., editors. Toxic Phytoplankton Blooms in the Sea. Elsevier; New York: 1993. p. 67-72.

Tester, PA.; Stumpf, RP.; Steidinger, KA. Ocean color imagery: What is the minimum detection level for Gymnodinium breve blooms?. In: Reguera, B.; Blanco, J.; Fernandez, ML.; Wyatt, T., editors. Harmful algae; Proceedings of the VIII International Conference on Harmful Algae; Paris, France. Xunta de Galacia and Intergovernmental Oceanographic Commission of UNESCO; 1998.

Tomlinson MC, Stumpf RP, Ransibrahmanakul V, Truby EW, Kirkpatrick GJ, Pederson BA, Vargo GA, Heil CA. Evaluation of the use of SeaWiFS imagery for detecting Karenia brevis harmful algal blooms in the eastern Gulf of Mexico. Remote Sens. Environ 2004;91:293–303.

Tomlinson MC, Wynne TT, Stumpf RP. An evaluation of remote sensing techniques for enhanced detection of the toxic dinoflagellate, Karenia brevis. Remote Sens. Environ 2009;113:598–609.

Walsh JJ, Steidinger KA. Saharan dust and Florida red tides: The cyanophyte connection. J. Geophys. Res 2001;106(C6):11,597–11,612.

Watkins SM, Reich A, Fleming LE, Hammond R. Neurotoxic Shellfish Poisoning. Mar. Drugs 2008;6 (3):431–455. [PubMed: 19005578]

Werdell J, Bailey S, Fargion G, Pietras C, Knobelspiesse K, Feldman G, McClain C. Unique Data Repository Facilitates Ocean Color Satellite Validation. EOS, Trans. Amer. Geophys. Uni 2003;84 (38):377–392.

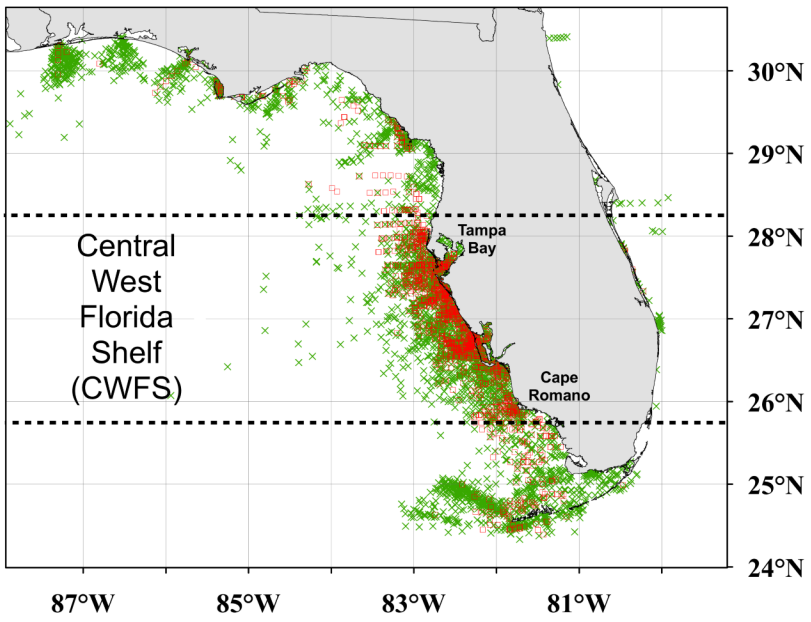

Figure 1.

Florida Harmful Algal Bloom Historical Database (2002 to 2006; data provided by the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute; Haverkamp et al., 2004). "Identified in situ samples to be considered" (see Figure 2) illustrating the sampling location for Karenia brevis: red tides ($\geq 1.5\times10^{4}$ cells l^{-1}; "□") and non-blooms ($<1.5\times10^{4}$ cells l^{-1}; "×"). The region of the Central West Florida Shelf (CWFS – south of 28.25°N close to Tampa Bay and north of 25.75°N near Cape Romano) is indicated.

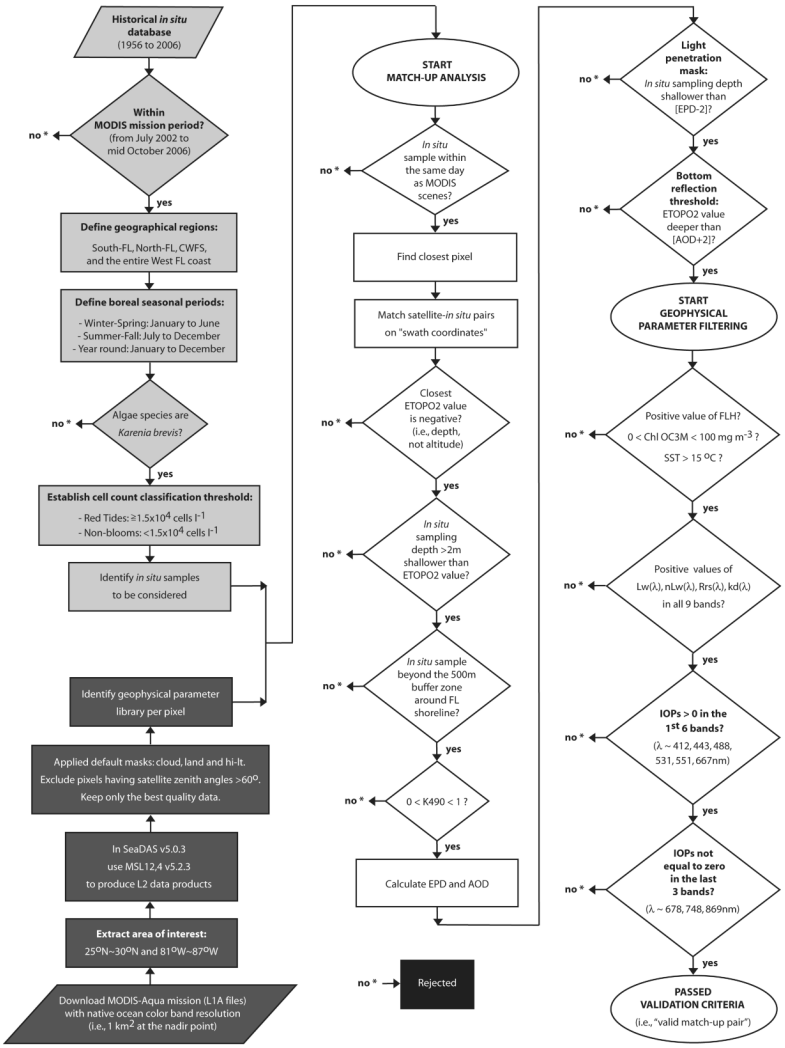

Figure 2.

Flowchart representing the stringent quality control tests applied to all near-coincident satellite-in situ pairs; adapted from Carvalho (2008) and Carvalho et al. (2010). Steps to identify in situ samples to be considered (light gray); satellite processing to identify the geophysical parameter library (dark gray); level-by-level of the match-up analysis performed to derive the valid cloud-free match-ups (white); and data removed from the analysis (black). As the water column depth was not specified for the in situ sampling within the Florida HAB Historical Database provided by the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute (Haverkamp et al., 2004), the Earth TOPOgraphy database at two-minute gridded resolution (ETOPO2 – http://www.ngdc.noaa.gov/mgg/gdas/gd_designagrid.html#) was used. Refer to Figure 6.3 in Carvalho (2008) for graphical representation of effective penetration depth (EPD; Gordon and McCluney; 1975) and apparent optical depth (AOD; Bailey and Werdell, 2006). Inherent optical properties (IOPs) includes: $a(\lambda)$, $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, $b(\lambda)$, $b_{bp}(\lambda)$. For additional and information see Section 2.

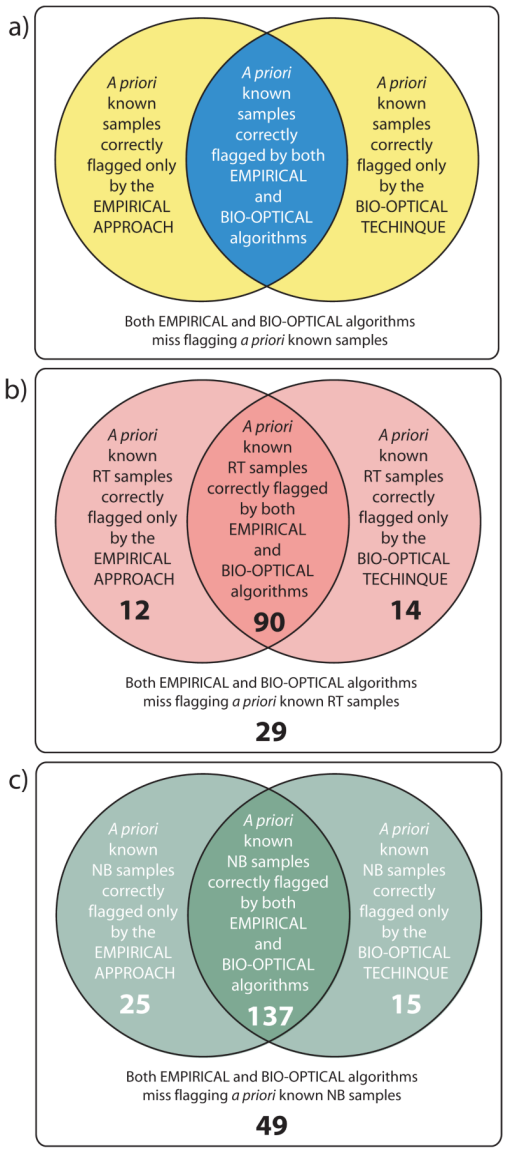

Figure 3.

Venn diagrams showing the structure of the two configurations of the Hybrid Scheme (Section 4), which is based on the detection frequency of the Empirical Approach and the Bio-optical Technique. a) The positive identifications of the 1st configuration are represented by YELLOW and BLUE (when either or both algorithms correctly flagged known samples), while the positive identifications of the 2nd configuration are represented only by the BLUE (both algorithms had to flag simultaneously known samples correctly). b, c) All valid match-ups (n=371=[145+226]) sampled along the Central West Florida Shelf during the Summer-Fall periods (2002 to 2006). b) Shown in RED are red tide samples (RT; $\geq 1.5\times10^4$ cells l^{-1}; n=145=[12+90+14+29]). c) Shown in GREEN are non-bloom samples (NB; $<1.5\times10^4$ cells l^{-1}; n=226=[25+137+15+49]). See Tables 1 and 2 for the detection frequency of the two configurations of the Hybrid Scheme.

Table 1

Detection frequency of the Hybrid Scheme along the Central West Florida Shelf during the Summer-Fall periods (2002 to 2006). Panels I-II: Error matrices (use Table 2 for assistance with A, B, C, and D); Panels IV-VI: 1st configuration of the Hybrid Scheme (when either or both algorithms needed to flag known samples correctly; see Figure 3); Panels VII-IX: 2nd configuration of the Hybrid Scheme (when both algorithms must have flagged simultaneously known samples correctly; see Figure 3); Panels X-XII: Individual versions of the Empirical Approach and the Bio-optimal Technique, respectively (* adapted from Table 6 in Carvalho et al. (2010)). Panels II, V, VIII, and XI take into account the rows of the table and are based on the number of a priori known in situ samples (i.e., "in situ cell count"), where the RED indicates red tide samples ($\geq 1.5\times10^4$ cells l^{-1}) and the GREEN indicates non-blooming conditions ($<1.5\times10^4$ cells l^{-1}); these panels answer the first set of questions (Section 3: 1a-1d): sensitivity (Sens; A/[A+B]; Question 1a), false negative (Fneg; B/[A+B]; Question 1b), false positive (Fpos; C/[C+D]; Question 1c), and specificity (Spec; D/[C+D]; Question 1d). Panels III, VI, IX, and XII take into account the columns and are based on the identifications in the satellite data (i.e., "algorithm outcome"), where the RED indicates samples flagged as HAB-positive and the GREEN indicates samples classified as HAB-negative; these panels answer the second set of questions (Section 3: 2a-2d): positive predictive value (posPV; A/[A+C]; Question 2a), inverse of the positive predictive value (IposPV; C/[A+C]; Question 2b), inverse of the negative predictive value (InegPV; B/[B+D]; Question 2c), negative predictive value (negPV; D/[B+D]; Question 2d). For an ideal algorithm, B=C=0.

I) Detection frequency

1st configuration of the Hybrid Scheme		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)	Sum
In situ cell count	True (RT)	A	B	A+B
	False (NB)	C	D	C+D
	Sum	A+C	B+D	total

II) 1st set of questions (1a-1d)

1st configuration of the Hybrid Scheme		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)	Sum
In situ cell count	True (RT)	A/[A+B] Sens	B/[A+B] Fneg	100%
	False (NB)	C/[C+D] Fpos	D/[C+D] Spec	100%

III) 2nd set of questions (2a-2d)

1st configuration of the Hybrid Scheme		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)
In situ cell count	True (RT)	A/[A+C] posPV	B/[B+D] InegPV
	False (NB)	C/[A+C] IposPV	D/[B+D] negPV
	Sum	100%	100%

IV) Detection frequency

1st configuration of the Hybrid Scheme		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)	Sum
In situ cell count	True (RT)	116	29	145
	False (NB)	49	177	226
	Sum	165	206	371

V) 1st set of questions (1a-1d)

1st configuration of the Hybrid Scheme		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)	Sum
In situ cell count	True (RT)	80% Sens	20% Fneg	100%
	False (NB)	22% Fpos	78% Spec	100%

VI) 2nd set of questions (2a-2d)

1st configuration of the Hybrid Scheme		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)
In situ cell count	True (RT)	70% posPV	14% InegPV
	False (NB)	30% IposPV	86% negPV
	Sum	100%	100%

VII) Detection frequency

2nd configuration of the Hybrid Scheme		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)	Sum
In situ cell count	True (RT)	90	55	145
	False (NB)	89	137	226
	Sum	179	192	371

VIII) 1st set of questions (1a-1d)

2nd configuration of the Hybrid Scheme		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)	Sum
In situ cell count	True (RT)	62% Sens	38% Fneg	100%
	False (NB)	39% Fpos	61% Spec	100%

IX) 2nd set of questions (2a-2d)

2nd configuration of the Hybrid Scheme		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)
In situ cell count	True (RT)	50% posPV	29% InegPV
	False (NB)	50% IposPV	71% negPV
	Sum	100%	100%

X) Detection frequency

Individual optimized algorithms *		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)	Sum
In situ cell count	True (RT)	102, 104	43, 41	145
	False (NB)	74, 64	152, 162	226
	Sum	176, 168	195, 203	371

XI) 1st set of questions (1a-1d)

Individual optimized algorithms *		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)	Sum
In situ cell count	True (RT)	~70% Sens	~30% Fneg	100%
	False (NB)	~30% Fpos	~70% Spec	100%

XII) 2nd set of questions (2a-2d)

Individual optimized algorithms *		Algorithm outcome: True (RT)	Algorithm outcome: False (NB)
In situ cell count	True (RT)	~60% posPV	~20% InegPV
	False (NB)	~40% IposPV	~80% negPV
	Sum	100%	100%

Table 2

Detection frequency of the Hybrid Scheme: 1st configuration (Left) and 2nd configuration (Right). The RED shows the incidence of red tides (n=[A+B]; samples containing ≥1.5 x10^4 cells l^{-1}), and the GREEN shows non-blooming conditions (n=[C+D]; samples containing <1.5x10^4 cells l^{-1}). Use error matrices in Table 1 [panels I, IV, and VII] and Venn diagrams in Figure 3 to assist with the interpretation and for the use of A, B, C and D to access the accuracy of the Hybrid Scheme.

1st configuration of the Hybrid Scheme		RED TIDE SAMPLES	2nd configuration of the Hybrid Scheme	
	90	***A priori* known samples correctly flagged** BY BOTH ALGORITHMS	90	A = 90
A = 116	12	***A priori* known samples correctly flagged** ONLY BY THE EMPIRICAL APPROACH	12	
	14	***A priori* known samples correctly flagged** ONLY BY THE BIO-OPTICAL TECHNIQUE	14	B = 55
B = 29	29	BOTH ALGORITHMS **miss flagging samples**	29	
A + B	145	Total red tide population	145	A + B

1st configuration of the Hybrid Scheme		NON-BLOOM SAMPLES	2nd configuration of the Hybrid Scheme	
C = 49	49	BOTH ALGORITHMS **miss flagging samples**	49	
	25	***A priori* known samples correctly flagged** ONLY BY THE EMPIRICAL APPROACH	25	C = 89
D = 177	15	***A priori* known samples correctly flagged** ONLY BY THE BIO-OPTICAL TECHNIQUE	15	
	137	***A priori* known samples correctly flagged** BY BOTH ALGORITHMS	137	D = 137
C + D	226	Total non-bloom population	226	C + D

1st configuration of the Hybrid Scheme	STATISTICAL ASSESSMENT	2nd configuration of the Hybrid Scheme
([116+29]/371)~39%	Prevalence of red tide = ([A+B]/[A+B+C+D])	([90+55]/371)~39%
([116+177]/371)~79%	Overall Accuracy = ([C+D]/[A+B+C+D])	([90+137]/371)~62%

A = *A priori* known red tide samples correctly flagged as HAB-positive.

B = *A priori* known red tide samples incorrectly flagged as HAB-negative

C = *A priori* known non-bloom samples incorrectly flagged as HAB-positive

D = *A priori* known non-bloom samples correctly flagged as HAB-negative