Satellite Retrievals of Karenia brevis Harmful Algal Blooms in the West Florida Shelf Using Neural Networks and Comparisons with Other Techniques

Ahmed El-habashi
CUNY City College

Ioannis Ioannou
NATO Science & Technology Organization

Michelle Tomlinson
NOAA National Centers for Coastal Ocean Science

Richard P. Stumpf
NOAA National Centers for Coastal Ocean Science

Sam Ahmed
CUNY City College

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/cc_pubs

Part of the Electrical and Computer Engineering Commons, Ocean Engineering Commons, and the Oceanography and Atmospheric Sciences and Meteorology Commons

Recommended Citation
El-habashi, Ahmed; Ioannou, Ioannis; Tomlinson, Michelle; Stumpf, Richard P.; and Ahmed, Sam, "Satellite Retrievals of Karenia brevis Harmful Algal Blooms in the West Florida Shelf Using Neural Networks and Comparisons with Other Techniques" (2016). CUNY Academic Works.
https://academicworks.cuny.edu/cc_pubs/449

This Article is brought to you for free and open access by the City College of New York at CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.
Article

Satellite Retrievals of Karenia brevis Harmful Algal Blooms in the West Florida Shelf Using Neural Networks and Comparisons with Other Techniques

Ahmed El-habashi 1,†, Ioannis Ioannou 2,†, Michelle C. Tomlinson 3,†, Richard P. Stumpf 3,† and Sam Ahmed 1,*†

1 The City College of New York, Optical Remote Sensing Laboratory, Department of Electrical Engineering, 160 Convent Ave, New York, NY 10031, USA; aelhaba00@citymail.cuny.edu
2 NATO Science & Technology Organization, Center for Maritime Research and Experimentation, Viale San Bartolomeo 400, La Spezia 19126, Italy; yannismail@gmail.com
3 NOAA National Centers for Coastal Ocean Science, 1305 East-West Highway Code N/SCI1 Silver Spring, MD 20910, USA; michelle.tomlinson@noaa.gov (M.C.T.); richard.stumpf@noaa.gov (R.P.S.)
* Correspondence: ahmed@ccny.cuny.edu; Tel.: +1-212-650-7250; Fax: +1-212-650-8249
† These authors contributed equally to this work.

Academic Editors: Deepak R. Mishra, Richard W. Gould Jr., Xiaofeng Li and Prasad S. Thenkabail

Received: 30 December 2015; Accepted: 20 April 2016; Published: 4 May 2016

Abstract: We describe the application of a Neural Network (NN) previously developed by us, to the detection and tracking, of Karenia brevis Harmful Algal Blooms (KB HABs) that plague the coasts of the West Florida Shelf (WFS) using Visible Infrared Imaging Radiometer Suite (VIIRS) satellite observations. Previous approaches for the detection of KB HABs in the WFS primarily used observations from the Moderate Resolution Imaging Spectroradiometer Aqua (MODIS-A) satellite. They depended on the remote sensing reflectance signal at the 678 nm chlorophyll fluorescence band (Rrs678) needed for both the normalized fluorescence height (nFLH) and Red Band Difference (RBD) algorithms currently used. VIIRS which has replaced MODIS-A, unfortunately does not have a 678 nm fluorescence channel so we customized the NN approach to retrieve phytoplankton absorption at 443 nm (a\text{\text{ph}}443) using only Rrs measurements from existing VIIRS channels at 486, 551 and 671 nm. The a\text{\text{ph}}443 values in these retrieved VIIRS images, can in turn be correlated to chlorophyll-a concentrations [Chla] and KB cell counts. To retrieve KB values, the VIIRS NN retrieved a\text{\text{ph}}443 images are filtered by applying limiting constraints, defined by (i) low backscatter at Rrs 551 nm and (ii) a minimum a\text{\text{ph}}443 value known to be associated with KB HABs in the WFS. The resulting filtered residual images, are then used to delineate and quantify the existing KB HABs. Comparisons with KB HABs satellite retrievals obtained using other techniques, including nFLH, as well as with in situ measurements reported over a four year period, confirm the viability of the NN technique, when combined with the filtering constraints devised, for effective detection of KB HABs.

Keywords: neural networks; harmful algal blooms; ocean color remote sensing reflectance; Karenia brevis; retrieved chlorophyll-a; normalized fluorescence height; West Florida Shelf

1. Introduction

A new approach is needed 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. Such a monitoring capability for KB HABs is important because of their negative impacts on ecology and health. More specifically, high KB HABs levels pose a threat to fisheries and human health, and directly affect tourism and local economies [1]. Effective KB HABs detection and tracking approaches are needed for use with VIIRS so that NOAA can extend its HABs monitoring capabilities, which previously relied on...
MODIS-A imagery [2–11]. Unfortunately, VIIRS, unlike its predecessor MODIS-A, does not have a 678 nm channel to detect chlorophyll fluorescence, which is used in the normalized fluorescence height (nFLH) algorithm [7,8], or in the Red Band Difference (RBD) algorithm, developed by our CCNY group [5,6]. Both of these techniques have demonstrated that the Rs 678 nm from the MODIS-A fluorescence band helps in effectively detecting and tracking KB HABs in the WFS [3–8]. To overcome the lack of a fluorescence channel on VIIRS, the alternative approach described here, bypasses the need for measurements of chlorophyll fluorescence, allowing us to extend KB HABs satellite monitoring capabilities in the WFS to VIIRS.

The essence of the approach is the application of a standard multiband neural network (NN) inversion algorithm, previously developed and reported by us [12–15], that takes VIIRS Rs measurements at the 486, 551 and 671 nm bands (or 488, 555, and 667 nm for MODIS-A) as inputs, and produces as output the related Inherent Optical Properties (IOPs), namely: absorption coefficients of phytoplankton (a_{443}) dissolved organic matter (a_g) and non-algal particulates (a_{dm}) as well as the particulate backscatter coefficient, (b_{bp}) all at 443 nm. In this work, however, it is only with the NN output of a_{443} that we are concerned with and which can be converted to equivalent [Chla] using empirical relationships for specific absorption in the WFS which have been determined from in situ measurements [16]. For the next step, to obtain KB values from the VIIRS NN retrieved a_{443} image, we apply additional constraints, defined by (i) low backscatter manifested as a maximum Rs551 value and (ii) a minimum [Chla] threshold [4,17,18] (and hence an equivalent minimum a_{443}min value) both known to be associated with KB HABs in the WFS. These two constraining filter processes are applied sequentially to the VIIRS NN retrieved a_{443} image. First an image is made of retrieved VIIRS Rs551. A mask is then made of all pixels with Rs551 \geq Rs551max, the maximum value known to be compatible with the existence KB HABs. This mask is then applied, in a filter process hereinafter dubbed F1, to the VIIRS NN retrieved a_{443} image to exclude pixels with Rs551 \geq Rs551max. The residual image, after application of filter process F1, will only show a_{443} values that also comply with Rs551 \leq Rs551max. Then, in a second filter process, hereinafter dubbed F2, all values of a_{443} \leq a_{443}min are eliminated. The residual image will now only show a_{443} values that are compatible with both criteria for KB HABs, and are therefore representative of KB HABs.

It will be shown that when both these filter condition are applied to VIIRS NN a_{443} retrievals, they can be used to effectively delineate and quantify KB HABs in the WFS. The KB HABs retrieved in this manner also show good correlations with in situ KB HABs measurements as well as with nFLH retrievals and other techniques to which the same filtering criteria have been applied, confirming the viability of the approach.

2. Materials and Methods

Neural Network Algorithm Background

In the work presented here, we use our synthetically trained NN algorithm [12–15] to solve the inverse problem [19,20] of retrieving physical variables, including a_{443}, from VIIRS observations of Rs at 486, 551 and 671 nm in the WFS. A brief background summary is given here. Additional details are available in the cited references and in the Appendix. The latter also gives the necessary parameters for using the NN with a MATLAB tool for satellite retrievals by any interested reader. As mentioned in the introduction, we use a standard multiband NN inversion algorithm that takes VIIRS Rs measurements at the 486, 551 and 671 nm bands as inputs, and produces as outputs the related IOPs, namely: a_{443}, a_g and a_{dm} as well as b_{bp}, all at 443 nm. (As mentioned earlier, it is only with the a_{443} output of NN that we are concerned with here).

Synthetic Dataset

For development of the algorithm, available measurement data from the NASA Bio-Optical Marine Algorithm Data set [21], (NOMAD) was used as a basis for establishing a range of possible
IOP parameters and examining potential relationships between them. The full range of parameters, applicable to both oceanic and coastal waters, and the relationships established between them were used as inputs to a bio-optical model [14,22,23]. The range and variability of IOPs used in our bio-optical modeling is well represented in the literature [22–35]. Our model, as illustrated in Figure 1a, has four components: water, phytoplankton, colored dissolved organic matter (CDOM) and non-algal particulates (NAP) as inputs to create a synthetic data base consisting of 20,000 IOP combinations covering the entire observable range of possible IOP parameters. These IOPs were then used with a HydroLight based [36], parameterized forward model, described in Lee 2002 [23] to generate corresponding sets of \( R_{rs} \) values at 486, 551 and 671 nm for VIIRS, and at 488, 555 and 667 nm for MODIS-A, associated with each of the 20,000 IOPs (and the related different combinations of the four components) [12–15].

Algorithm Training

Once the comprehensive set of \( R_{rs} \) spectral channel data together with the associated IOP’s are assembled, a robust NN can be developed to connect the \( R_{rs} \) directly to the physical parameter to be retrieved. As is standard, a subset of the data is used to train the underlying NN and the rest used for validations. In our case, 50% (10,000) of these sets of \( R_{rs} \) and related IOPs were used [12–15] to train the NN model to obtain the relationship between input \( R_{rs} \) values at the 486, 551 and 671 nm bands for VIIRS and the output \( a_{ph}, a_{g}, a_{dm}, b_{hp} \) coefficients at 443 nm, which is at the peak of phytoplankton absorption spectrum, and thus exhibits the most sensitive variation in natural waters (and where \( a_g \) and \( a_d \) are mutually constrained through an empirically derived relationship based on a study of NOMAD [21] values). Figure 1b shows the VIIRS NN Architecture used. In our implementation, it was found that a one-hidden-layer multilayer perceptron (MLP-NN) with six neurons at the hidden layer was able to model our problems well. Since this architecture appeared appropriate for our needs, we did not further explore how this relationship might be improved by increasing the number of neurons or layers.

The VIIRS NN algorithm was tested and evaluated using multiple experiments. It was (a) tested initially on half the simulated data set that was not used in the training (b) field data from the,
(NOMAD) and (c) data from our own field campaign in the Chesapeake Bay during 2013, which represents well the typical range of water optical properties and chlorophyll-a concentrations in coastal regions. These tests confirmed good retrievals of \( a_{ph,443} \) by the NN [13,14] and provided strong motivation for its use in the current work. The results of the present work have further confirmed its viability for effective retrievals of \( a_{ph,443} \) in the coastal waters of the WFS from VIIRS measurements at the 486, 551 and 671 nm bands. In this context, we would also add that though we are estimating \( a_{ph,443} \) at 443 nm, the \( R_{rs} \) values used as inputs to the NN to retrieve it are at longer wavelengths, 486, 551 and 671 nm, and do not require measurements at the shorter wavelength channels, M1 (415 nm) and M2 (443 nm) which are typically most impacted by atmospheric correction difficulties.

As mentioned above, a more detailed description of the NN, and its background are given in the Appendix. The parameters needed for its implementation and the retrieval \( a_{ph,443} \) from inputs of \( R_{rs} \) 486, 551 and 671 nm. The related calculations that can be readily performed by interested readers using a MATLAB script are also described in the Appendix. We would be interested to hear the results of other retrievals, particularly for other areas where KB HABs are known to exist.

3. Results and Discussion

3.1. VIIRS Retrievals of \( R_{rs,551} \) and \( a_{ph,443} \) and Determination of Limiting Values of \( R_{rs,551}^{\text{max}} \) and \( a_{ph,443}^{\text{min}} \) in a KB Bloom Environment

Figure 2 shows KB cell counts associated with a prominent KB HAB bloom in the WFS with its peak occurring around 2 September 2014. This bloom has been extensively studied [7,8,15]. Figure 2, from the NOAA HABSOS site [37] shows the bloom area with indicators (circles of different sizes) for KB HABs cell counts for concurrent or near concurrent dates over the period (27 August–17 September 2014). Figure 3a shows VIIRS retrievals of \( R_{rs,551} \) for 2 September 2014 and for the same location. It shows, overlaid, the cell count information from Figure 2. Considering backscatter data as a KB indicator, Refs. [3,17,18] give maximum permissible values of backscatter, compatible with KB HABs, as \( bb_{p,550} \leq 0.0045 \text{ m}^2 \cdot \text{mg}^{-1} \) at 550 nm. In addition, for \( R_{rs} \) 551 which is taken to serve as a proxy for backscatter, the equivalent max permissible value is \( R_{rs,551} \leq 7.0 \times 10^{-3} \text{ sr}^{-1} \). However, by inspection of Figure 3a of retrieved [38] and \textit{in situ} [37] values we concluded that a value of approximately \( R_{rs,551} \leq 6.0 \times 10^{-3} \text{ sr}^{-1} \) appears more appropriate as the highest permissible limiting value for compatibility with KB HABs. If that value is applied to Figure 3a, as filter process F1, all pixels with \( R_{rs,551} \geq 6.0 \times 10^{-3} \text{ sr}^{-1} \) are screened out, leaving the residual pixels in Figure 3b as compatible with KB HABs as far as backscatter is concerned.

![Figure 2. NOAA HABSOS data [37] with \textit{in situ} KB concentrations, for period 8 August–17 September 2014.](image)

**Cell Counts/L Classification:**
- Not Observed
- Very Low (1-10,000)
- Low (10,000-100,000)
- Medium (100,000-1,000,000)
- High (1,000,000+)
We next apply the pixel mask shown in Figure 3b, resulting from application of Filter 1 conditions to VIIRS Rs551 retrievals, to Figure 4a which shows \(a_{\text{ph}443}\) retrieved by the NN for the same 2 September 2014 date and location. This process eliminates all \(a_{\text{ph}443}\) values in Figure 4a that do not satisfy \(Rs551 \leq 6.0 \times 10^{-3} \text{ sr}^{-1}\) requirements for compatibility with KB HABs. The residual values, shown in Figure 3b satisfy the Rs551 limit requirements. Cell counts from Figure 2 are again shown overlaid in Figure 4b.

We next examine the appropriate F2 filter values for minimum \(a_{\text{ph}443}\) consistent with KB HABs. Now Refs. [17,18] give the minimum \([\text{Chla}]\) concentration compatible with KB HABs in the WFS as \([\text{Chla}] \geq 1.5 \mu\text{g} \cdot \text{L}^{-1}\). This is converted into equivalent \(a_{\text{ph}443}\) using the empirical relationship [16]:

\[
a_{\text{ph}443} = 0.051 \times [\text{Chla}]^{0.74} \text{ m}^{-1} \tag{1}
\]

Which gives \(a_{\text{ph}443} \geq 0.0688 \text{ m}^{-3}\) as the minimum permissible value. Again, by inspection of Figure 4b against overlaid in situ values, we conclude that a value of approximately \(a_{\text{ph}443} \geq 0.061 \text{ m}^{-3}\) is more appropriate as the limiting value for compatibility with KB HABs.

Applying this value as Filter F2 to Figure 4b we eliminate pixels that are not compatible with KB and end up with Figure 5a. This shows only residual values with \(a_{\text{ph}443} \geq 0.061 \text{ m}^{-3}\) which also satisfy \(Rs551 < 6.0 \times 10^{-3} \text{ sr}^{-1}\). These are therefore residual values compatible with and representative of KB HABs. Comparison of the residual \(a_{\text{ph}443}\) (left hand scale) and equivalent \([\text{Chla}]\) (right hand scale) in Figure 5a, which represent the retrieved KB HABs values, show good qualitative agreement with the KB HABs cell counts in Figure 2.
We next compare the NN retrievals of KB HABs, described in the last section, with retrievals using other techniques. In doing so, it is not the intention to present an exhaustive comparison of the many approaches that have been studied for the detection of KB HABs in the WFS. Some of these are listed in [2–11]. For more details, the reader is referred to recent excellent surveys and comparisons in [3,11]. We focus here, instead, on comparisons with three distinct categories of retrieval techniques.
The existing satellite algorithms that are relevant to the detection of KB HABs in the WFS rely on retrieved chlorophyll-a concentrations [Chla] as part of their processing. They retrieve it primarily in one of three ways: (i) using the remote sensing reflectance fluorescence signal at Rrs678 measured by the MODIS-A satellite; and (ii) using blue-green ratio algorithms, exemplified by the NASA OC3 and OCI products [21,39–41]; and (iii) a more recent WFS region specific empirical algorithm, the Red Green chlorophyll-a Index (RGCI) which retrieves [Chla] using 2 visible VIIRS or MODIS-A bands [42]. Those methods (i) using the 678 nm fluorescence signal, include the normalized fluorescence height (nFLH) [2–11] and Red Band Difference (RBD) techniques [5,6]. (nFLH is FLH normalized to solar radiance); Those methods (ii) using blue-green ratio algorithms, include the chlorophyll-a anomaly technique and related approaches [43–52]. The method using RGCI, has been modified very recently, and has been reported for [Chla] retrievals but has not yet been applied to KB retrievals [42].

Amongst the various available KB HABs retrieval techniques that have been comprehensively evaluated, nFLH and RBD exhibited the best KB HABs retrieval statistics [3] for the WFS. Accordingly, our paper focuses on comparisons of NN retrievals against retrievals using (i) nFLH/RBD fluorescence and (ii) OCI/OC3 approaches; and (iii) RGCI approaches. These are each described in Sections 3.2.1–3.2.4 respectively, with detailed descriptions available in references [2–11,42–53].

3.2.1. Comparison of NN KB HABs Retrieval Techniques with Those Using the Normalized Fluorescence Height (nFLH) and Red Band Difference (RBD) Techniques

Both nFLH and RBD techniques use the magnitude of Rrs at the 678 nm fluorescence band relative to the values at adjacent bands on MODIS-A to detect the chlorophyll fluorescence [54] and [Chla] [55]. This approach, when coupled with the known low backscatter properties of KB, has been shown to help in effectively detecting KB in the WFS [3,17,18].

Since the nFLH technique requires the fluorescence signal at 678 nm, which is absent from VIIRS, comparisons of nFLH with the NN technique must be made using MODIS-A neural network retrievals. For this purpose, the same approach used with VIIRS NN (above) is used to obtain MODIS-A NN a ph443 retrievals. For MODIS-A, the Rrs values at the 488, 555 and 667 nm bands are used as inputs to the NN. For application of the F1 filter process, Rs 555 nm is used as the proxy for backscatter (instead of the Rs 551 nm used with VIIRS). The same F1 and F2 filter processes, with the same limiting values of Rs 555 ≤ 6.0 × 10^{-3} sr^{-1} for F1, and a ph443 ≥ 0.061 m^{-3} for F2, were applied to MODIS-A NN retrievals for the same 2 September 2014 bloom date, to obtain KB compatible residual retrieval values. Figure 5b shows the MODIS-A NN retrieved a ph443 (left hand scale) and the equivalent [Chla] (right hand scale), after application of F1 and F2 filters. As can be seen, Figure 5b shows good qualitative agreement with the images of Figure 5a,b.

We next compare the nFLH NASA product available [54] with MODIS-A for retrievals of nFLH values for the same 2 September 2014 WFS bloom. nFLH uses the height of the Rrs 678 nm (over a baseline connecting adjacent Rrs values) as an indicator of chlorophyll fluorescence magnitude [54,55]. Retrieved values are shown in Figure 6a. The left hand scale shows the nFLH NASA product, while the right hand scale shows the equivalent [Chla] value, converted using the empirical relationship [7]:

\[
[Chla] = 1.255 \times (FLH \times 10)^{0.86}, \mu g L^{-1}
\]  

(2)

Figure 6b shows the same nFLH retrievals, but with filters F1 and F2 applied, i.e., with the same pixels filtered out as in the MODIS-A NN retrievals in Figure 5b. It can be seen that Figure 6b shows good qualitative agreement with the images of Figure 5a,b.

For more quantitative comparisons, Figure 7a shows the correlation between equivalent [Chla] derived from nFLH in Figure 6a, against the retrieved MODIS-A NN derived equivalent [Chla]. To avoid complications from more complex waters in-shore areas, the comparison is made for the off-shore area A, shown demarcated in Figure 6a,b above. Without the KB compatibility filters applied, a high correlation, R^2 = 0.87 is obtained. With the F1 and F2 filters applied (eliminating non KB compatible pixels) the correlation becomes R^2 = 0.71, Figure 7b. The color coding of the dots in
Figure 7a,b indicates distance from shore, with red being the closest. As can be seen the spread gets larger closer to the shore, where waters are more complex and the uncertainty increases.

![Figure 6](image_url)  
**Figure 6.** (a) MODIS-A nFLH (left hand scale) standard ocean color product, and equiv. [Chla] (right hand scale); (b) Same image as Figure 6a after filter process masks F1 and F2 are applied. Residual values are therefore compatible with and indicate KB blooms. Notes that “×” symbol in both images indicates no cell counts observed, in situ within 2 h windows from MODIS-A observation. Dark gray represents F1&F2 masks and white represent cloud cover or invalid data.

![Figure 7](image_url)  
**Figure 7.** (a) MODIS-A nFLH equiv. [Chla] retrievals against MODIS-A NN equiv. [Chla] retrievals; (b) Shows the same retrievals as 5a but with F1 and F2 masks applied. Color coding of the dots denotes distance to shore, with red being the closest.

When the comparison between nFLH and VIIRS NN retrieved [Chla] is expanded beyond area A to include all areas, including the more complex in-shore waters, the correlation between them drops drastically, to $R^2 = 0.19$, shown in Figure 8. This may be due to the poor retrievals associated with nFLH for in-shore areas, where increased backscatter can cause an apparent increase in the red-near infrared spectral peak, increasing the apparent nFLH chlorophyll fluorescence signal, but which is in
As mentioned above, RBD, along with nFLH, has exhibited the best statistics for agreement with both nFLH, MODIS-A NN on 2 September 2014 bloom, with filters F1 and F2 applied. As can be seen, this shows good qualitative confirmation of presence of KB blooms. The RBD index is defined by:

\[
RBD = nLw(678) - nLw(667), \text{W} \cdot \text{m}^{-2} \cdot \text{um} \cdot \text{sr}^{-1}
\]

Figure 8 shows the RBD index, as defined above, retrieved from MODIS-A observations for the 2 September 2014 bloom, with filters F1 and F2 applied. As can be seen, this shows good qualitative agreement with both nFLH, MODIS-A NN [Chla], and HABSOS \textit{in situ} images, Figures 2, 5b and 6b. As mentioned above, RBD, along with nFLH, has exhibited the best statistics for KB HABs retrievals [3].
3.2.3. Comparisons of VIIRS NN Retrievals with VIIRS Ocean Color Chlorophyll-α (OCI/OC3) Retrievals

We also compared the NN retrievals from VIIRS with retrievals using the OCI/OC3 indexes, both NASA products [53] for the same 2 September 2014 bloom. OCI and the related OC3 NASA products are [Chla] retrieval algorithms that make use of the ratio of blue/green bands in the MODIS-A and VIIRS satellites [21,39–41], and were found to yield exactly the same retrievals for our WFS conditions.

Correlations between VIIRS OCI/OC3 [Chla] and VIIRS NN equivalent [Chla] retrievals are shown in Figure 10a for the 2 September 2014 bloom for the off-shore area A. Without filter applications, $R^2 = 0.9$. When filters F1 and F2 are applied, the residual values consistent with KB HABs show, Figure 10b, a correlation of $R^2 = 0.78$. Figure 10c shows a comparison for area A of nFLH equivalent [Chla] retrievals against OCI/OC3 [Chla] retrievals with an $R^2 = 0.75$ without filters. Figure 10d shows $R^2 = 0.50$ with filters F1 and F2 applied, considerably lower than the correlation $R^2 = 0.71$, Figure 7b, between nFLH equivalent [Chla] versus NN equivalent [Chla].

![Figure 10](image-url)

**Figure 10.** (a) Retrieved VIIRS OCI/OC3 [Chla] against retrieved MODIS-A NN [Chla]; (b) Same retrieval as 10a with Filters F1 and F2 applied; (c) Retrieved OCI/OC3 [Chla] against MODIS-A nFLH equiv. [Chla]; (d) Same as 10c but with filter F1 & F2 applied. The vertical line shows the NN [Chla] KB threshold (F2 mask) consistent with the existence of KB, residual pixels of F1 and F2 masks satisfy both maximum backscatter and minimum NN [Chla] and represent values for KB blooms. Color coding of the dots denotes distance to shore, with red being the closest.

Again, when the comparison between OCI/OC3 and MODIS-A NN retrieved [Chla] is expanded beyond area A to include all areas, including the more complex in-shore waters, the correlation between...
them drops drastically, to $R^2 = 0.07$ (not shown) from the $R^2 = 0.75$ in Figure 10c for the off-shore area A. This may be due to the poorer retrievals associated with both nFLH and OCI/OC3 for more complex in-shore waters with increased backscatter [3,21,39–41,53,56].

3.2.4. Comparisons of VIIRS NN Retrievals with Red/Green Chlorophyll-a Index (RGCI) Retrievals

We next examine RGCI $[Chla]$ retrievals for the same 2 September 2014 date and locations as in the above Sections. Equation (3a) below gives the relationship for RGCI MODIS-A $[Chla]$ retrievals, and Equation (3b) that for RGCI VIIRS $[Chla]$ retrievals [42].

$$ [Chla]_{RGCI} = 0.1 \times \exp \left( 11.8 \times \frac{Rrs_{667}}{Rrs_{547}} \right), \mu g L^{-1} \quad (3a) $$

$$ [Chla]_{RGCI} = 0.1 \times \exp \left( 11.8 \times \frac{Rrs_{671}}{Rrs_{551}} \right), \mu g L^{-1} \quad (3b) $$

We first compare VIIRS RGCI $[Chla]$ retrievals against VIIRS NN $[Chla]$ retrievals. Figure 11a shows a correlation of $R^2 = 0.72$ for open ocean area A. This drops to $R^2 = 0.49$ when filters F1 and F2 are applied, Figure 11b. This compares with $R^2 = 0.9$ and $R^2 = 0.78$ for the equivalent comparisons between OCI/OC3 and VIIRS NN $[Chla]$.

Figure 11. (a) Retrieved VIIRS RGCI $[Chla]$ against retrieved VIIRS NN $[Chla]$; (b) Same retrieval as 11a with Filters F1 and F2 applied; (c) Retrieved RGCI $[Chla]$ against MODIS-A nFLH equiv. $[Chla]$; (d) Same as 11c but with filter F1 & F2 applied. The vertical line shows the NN $[Chla]$ KB threshold (F2 mask) Color coding of the dots denotes distance to shore, with red being the closest.
We next compare RGCI [Chla] against nFLH MODIS-A [Chla] retrievals. We see that for the same off shore area A, Figure 11c, $R^2 = 0.72$. This is lower than the $R^2 = 0.87$ obtained for the equivalent comparison of nFLH against MODIS-A NN [Chla] retrievals, Figure 7a, above. When the comparison is not restricted to the off-shore areas, the correlation for RGCI-nFLH drops off completely to $R^2 = 0.0$ (not shown) compared to the previously given equivalent correlations for nFLH-MODIS-A NN of $R^2 = 0.19$, Figure 8 and $R^2 = 0.07$ for OCI/OC3 (not shown).

When the F1 and F2 filtering procedures are applied to the RGCI [Chla] retrievals, for area A, the correlation against nFLH MODIS-A [Chla] retrievals drops to $R^2 = 0.50$, Figure 11d. This compares to $R^2 = 0.71$ and $R^2 = 0.50$ for the equivalent correlations, shown previously, between nFLH and MODIS-A NN and MODIS-A OCI/OC3 respectively.

The salient points of Sections 3.2.1–3.2.4 show that in open ocean area A, nFLH [Chla] retrievals from MODIS-A are well correlated with MODIS-A NN [Chla] retrievals $R^2 = 0.87$, less so with equivalent MODIS-A OCI/OC3 with $R^2 = 0.75$, and MODIS-A RGCI with $R^2 = 0.72$. All correlations drop when filters F1 and F2 are applied to obtain KB HABs retrievals to $R^2 = 0.71$, $R^2 = 0.5$ and $R^2 = 0.5$, respectively. For comparisons not restricted to the off-shore area A, all correlations with nFLH MODIS-A [Chla] retrievals drop drastically to $R^2 = 0.19$, $R^2 = 0.07$ and $R^2 = 0.0$ for MODIS-A NN, MODIS-A OCI/OC3 and RGCI respectively. It should also be noted that all four retrieval techniques compared show qualitatively similar KB HABs retrieval images of the type shown in Figures 5a,b, 6b, and 9 (not shown for OCI/OC3 and RGCI).

### 3.3. Comparisons of VIIRS NN, OCI/OC3 and RGCI Retrievals with in Situ Cell Count Measurements

#### 3.3.1. Evaluation of VIIRS KB HABs Retrievals against in Situ Measurement for Match-Ups Occurring over 2012–2015 Period

The ultimate test for the viability of KB HABs satellite retrieval techniques is their ability to match retrieved values with concurrent in situ measurements. It is, however, difficult to find on any one day sufficient matches between satellite observations between 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, Figure 12, between VIIRS measurements and in situ data at concurrent dates and over the 2012–2015 period for which there was available VIIRS data. The conditions for match-up were: (1) pixel centers were 0.3 miles or less from the in situ measurement location; pixels are excluded from the match-up comparison if they had been flagged for land, cloud, 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, and any pixels which has water leaving radiance spectra with negative values in any one of its wavelengths. Cell counts sample measurements also had to be at least 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 $\sim 10^5$ cells $\cdot L^{-1}$ [49].

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 showed 94 cases which satisfied the match-up conditions specified above. Figure 13a, above, shows the relationship between NN retrieved $a_{ph443}$ and the in situ KB cell measurements for these 94 cases. This shows a correlation with $R^2 = 0.32$. Converting the $a_{ph443}$ [16] to equivalent [Chla] for these same 94 cases, shows, in Figure 13b, the same correlation against KB cell counts, with $R^2 = 0.32$ for both these match-ups. Figure 13c,d, show the equivalent match-up for the same 94 points for retrieved OCI/OC3 and RGCI [Chla] against in situ KB cell counts, which shows lower correlations with $R^2 = 0.28$ and $R^2 = 0.17$ respectively.
Figure 12. Showing in situ locations of Karenia brevis cell counts cover the range 0.01–9.2·10^6 cells·L\(^{-1}\) collected by Florida Fish and Wildlife Conservation Commission (FWC). Zoomed areas illustrate the extent of the underline details of KB values available in VIIRS retrievals for the period of 2012–2015.

Figure 13. In situ observation within the same day of VIIRS image: (a) VIIRS NN retrieved \(a_{ph443}\) against in situ KB cell counts; (b) VIIRS NN equiv. [Chla] against in situ cell counts; (c) VIIRS OCI/OC3 retrieved [Chla] against in situ KB cell counts; (d) VIIRS RGCI retrieved [Chla] against in situ KB cell counts. Color coding of the dots denotes distance to shore, with blue being the closest.
Now, it is recognized that temporal variations due to bloom drift etc. between satellite observation and in situ measurement times can affect these results. Accordingly, the results were separated in accordance with observation time windows. Those occurring with less than 1 h, and 0.5 h between satellite and in situ observations are shown separately. Figure 14a (1 h), b (0.5 h) show the match-ups between VIIRS NN retrieved equivalent \([\text{Chla}]\) against KB cell counts. The equivalent values for OCI/OC3 and RGCI \([\text{Chla}]\) retrievals are shown in Figure 14c–f respectively. For the 1 h window, Figure 14a, the correlation for the NN retrievals increases from the daily \(R^2 = 0.32\) to \(R^2 = 0.69\). For OCI/OC3 retrievals, it increases from \(R^2 = 0.28\) to \(R^2 = 0.33\). For RGCI it increases from \(R^2 = 0.17\) to \(R^2 = 0.62\). For the 0.5 h window, the correlations increase to \(R^2 = 0.82, R^2 = 0.44\) and \(R^2 = 0.7\) respectively. With the NN being the highest, closely followed by RGCI, then OCI/OC3, though the average errors appear to be twice as high for RGCI as compared with NN.

**Figure 14.** (a–f) Retrieved NN equiv. \([\text{Chla}]\) and OCI/OC3 \([\text{Chla}]\) and RGCI \([\text{Chla}]\) against in situ cell counts for 1 h and 30 min observation time windows. Note that the vertical color bar is indicates distant (mi) from coastline with red being closest to shore.
Thus, it is seen that reducing the time window between satellite and in situ observations can generally significantly increase the correlation between VIIRS retrieved [Chla] and KB cell counts. More area specific studies need to be carried out before definitive conclusions can be drawn.

We also use the above match-ups to assess the validity of the F1 and F2 filter limits for Rs551 \(\leq 6.0 \times 10^{-3} \text{ sr}^{-1}\) and \(a_{ph443} \geq 0.061 \text{ m}^{-3}\) defined in Section 3.1 above. Figure 15 shows that for the NN retrieved \(a_{ph443}\) match-ups falling within the \(\frac{1}{2}\) h window, which were retrieved without any filter constraints being applied on Rs551 and \(a_{ph443}\) all but two of them fall within these limits. Further refinements on these limiting filter values as a means of retrieving KB HABs will require planned future statistical analysis of false positives and negatives in retrieval results against in situ measurements.

![Figure 15](image)

Figure 15. Shows the match-ups falling within the \(\frac{1}{2}\) h window, with both filter thresholds.

Table 1 below shows summary table of retrieval comparisons, in situ results and KB thresholds criteria applied.

<table>
<thead>
<tr>
<th>y-axis</th>
<th>x-axis</th>
<th>(R^2)</th>
<th>Slope &amp; Intercept</th>
<th>(\varepsilon)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCI/OC3 [Chla] (µg·L(^{-1}))</td>
<td>VIIRS</td>
<td>NN-[Chla] (µg·L(^{-1}))</td>
<td>0.90</td>
<td>(y = 0.96x + 0.22)</td>
<td>0.33</td>
</tr>
<tr>
<td>RGC/ [Chla] (µg·L(^{-1}))</td>
<td></td>
<td></td>
<td>0.72</td>
<td>(y = 1.50x - 0.22)</td>
<td>0.95</td>
</tr>
<tr>
<td>nFLH [Chla]-KB (W·m(^{-2})·µm(^{-1})·sr(^{-1}))</td>
<td>MODIS</td>
<td>OCI/OC3 [Chla] (µg·L(^{-1}))</td>
<td>0.71</td>
<td>(y = 0.80x + 0.36)</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RGC/ [Chla] (µg·L(^{-1}))</td>
<td>0.50</td>
<td>(y = 0.48x + 0.70)</td>
<td>0.57</td>
</tr>
<tr>
<td>NN [Chla] (µg·L(^{-1}))</td>
<td></td>
<td></td>
<td>0.50</td>
<td>(y = 0.55x + 0.98)</td>
<td>0.57</td>
</tr>
<tr>
<td>OCI/OC3 [Chla] (µg·L(^{-1}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGC/ [Chla] (µg·L(^{-1}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NN [Chla] (µg·L(^{-1}))</td>
<td>In situ</td>
<td>KB Cell Counts (cells·L(^{-1}))</td>
<td>(</td>
<td>1\ h \text{ window}</td>
<td>)</td>
</tr>
<tr>
<td>OCI/OC3 [Chla] (µg·L(^{-1}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGC/ [Chla] (µg·L(^{-1}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NN [Chla] (µg·L(^{-1}))</td>
<td></td>
<td>KB Cell Counts (cells·L(^{-1}))</td>
<td>(</td>
<td>30\ min \text{ window}</td>
<td>)</td>
</tr>
<tr>
<td>OCI/OC3 [Chla] (µg·L(^{-1}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGC/ [Chla] (µg·L(^{-1}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Criteria Applied in WFS for KB-HABs Retrievals

<table>
<thead>
<tr>
<th>Rs551</th>
<th>NN (a_{ph443})</th>
<th>Equivalent NN [Chla] Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\leq 0.006 \text{ sr}^{-1})</td>
<td>(\geq 0.061 \text{ m}^{-3})</td>
<td>(\geq 1.27 \mu g \text{ L}^{-1})</td>
</tr>
</tbody>
</table>
3.3.2. Evaluation of NN KB HABs Retrievals for Specific Bloom Events

We also compared the ability of NN KB HABs retrievals against in situ KB HABs measurements for specific algal blooms. We apply filter processes F1 and F2 above to VIIRS NN $a_{ph443}$ retrievals to arrive at KB compatible values, in Figures 16 and 17a,b for three bloom dates: 28 August 2014, 16 November 2014 and 9 October 2012. The residual $a_{ph443}$ (left hand scale) and equivalent [Chla] values (right hand scale) are those compatible with KB HABs. Figure 16 shows the retrievals for the 8 August 2014, bloom now overlaid with actual KB cell counts for that date only (not for a range of overlapping dates as was used previously). There are a total of 20 match-ups on that day. There appears to be one false positive, indicated by the arrow. This presents an ambiguous situation, since it is near the edge of the bloom and changes may have occurred in the time window between in situ and satellite measurement, and secondly, while there is no in situ bloom measured at the surface at that date, there was one below the surface at nearly the same location, bringing into play the possibility of sub-pixel variations, indistinguishable when pixel averaging takes place. Figure 17a shows a bloom on 16 November 2014, near the coast, with a total of 6 match-ups. There are no false positives or negatives in this retrieval. This is particularly interesting since it is close to the coast, and therefore in more complex waters. Figure 17b shows a bloom on 9 October 2012, with 12 match-ups. It covers both near and offshore areas and has no false positives or negatives.

![Figure 16. VIIRS-NN KB HABs retrievals for blooms date (28 August 2014), showing bloom compatible $a_{ph443}$ and equiv. [Chla] values. Notes image are overlaid with cell counts for this date. White areas represent cloud cover or invalid data.](image)
Comparisons of the NN technique were made first with the nFLH and RBD techniques, which, in this approach, 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, as detailed above, the additional limiting constraints are applied, in two filter processes, F1 and F2 to eliminate from that image all \( a_{ph443} \) pixels which are not compatible with the existence of KB HABs. These limiting criteria are defined by (i) low backscatter manifested as a maximum VIIRS \( R_{rs} \) \( 551_{max} \) value \([3,17,18]\) and (ii) the minimum \([Chla]\) and hence equivalent minimum \( \left( a_{ph443} \right) \) values that are known to be associated with KB HABs in the WFS. Thus, through the implementation of filter processes F1 and F2, all pixels in the VIIRS NN retrieved \( a_{ph443} \) image that are incompatible with KB HABs are eliminated. The residual image then shows only retrieved \( a_{ph443} \) values and their equivalent \([Chla]\) values that are consistent with the existence of KB HABs. The specific limiting values for maximum \( R_{rs} \) \( 551 \) and minimum \( a_{ph443} \) used in the filter processes were based on published data, reinforced and refined by inspections of retrieval results against available \textit{in situ} data. The retrieved values of \( a_{ph443} \) and equivalent \([Chla]\) retrieved using these filter limits, were then compared both qualitatively and quantitatively both against other satellite retrieval techniques and \textit{in situ} measurements.

Comparisons of the NN technique were made first with the nFLH and RBD techniques, which, in reported comprehensive tests \([3]\) exhibited the best statistics for KB HABs retrievals in the WFS from open ocean waters. However, KB retrieval capabilities with the nFLH technique are severely limited.

4. Summary and Conclusions

Detection of KB HABs in the WFS have been aided in the past by use of the 678 nm MODIS-A fluorescence band for retrievals using nFLH and RBD techniques. The lack of a fluorescence channel on the successor VIIRS satellite provided the impetus for the NN work reported here. In this approach, 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, as detailed above, the additional limiting constraints are applied, in two filter processes, F1 and F2 to eliminate from that image all \( a_{ph443} \) pixels which are not compatible with the existence of KB HABs. These limiting criteria are defined by (i) low backscatter manifested as a maximum VIIRS \( R_{rs} \) \( 551_{max} \) value \([3,17,18]\) and (ii) the minimum \([Chla]\) and hence equivalent minimum \( \left( a_{ph443} \right) \) values that are known to be associated with KB HABs in the WFS. Thus, through the implementation of filter processes F1 and F2, all pixels in the VIIRS NN retrieved \( a_{ph443} \) image that are incompatible with KB HABs are eliminated. The residual image then shows only retrieved \( a_{ph443} \) values and their equivalent \([Chla]\) values that are consistent with the existence of KB HABs. The specific limiting values for maximum \( R_{rs} \) \( 551 \) and minimum \( a_{ph443} \) used in the filter processes were based on published data, reinforced and refined by inspections of retrieval results against available \textit{in situ} data. The retrieved values of \( a_{ph443} \) and equivalent \([Chla]\) retrieved using these filter limits, were then compared both qualitatively and quantitatively both against other satellite retrieval techniques and \textit{in situ} measurements.

Comparisons of the NN technique were made first with the nFLH and RBD techniques, which, in reported comprehensive tests \([3]\) exhibited the best statistics for KB HABs retrievals in the WFS from open ocean waters. However, KB retrieval capabilities with the nFLH technique are severely limited.

Figure 17. VIIRS-NN KB HABs retrievals on 2 different blooms dates, showing bloom compatible \( a_{ph443} \) and equiv. \([Chla]\) values. (a) 11 November 2014, bloom; (b) 09 October 2012, bloom. Notes all images are overlaid with cell counts corresponding for these dates. White areas represent cloud cover or invalid data.
MODIS-A. Our comparisons showed good correlations between nFLH and the NN retrievals for open ocean waters. However, KB retrieval capabilities with the nFLH technique are severely limited for in-shore retrievals, where the more complex waters with increased non-algal particulate scattering interfere with the fluorescence signal making its interpretation more difficult [56]. Comparisons were also made between the NN technique and retrievals using OCI/OC3 and the RGCI for [Chla] retrievals. Again, these comparisons showed reasonably correlations in open ocean waters, but confirmed difficulties with retrievals in more complex inshore waters. Further work is also needed on our NN KB retrieval technique before we can properly evaluate its capabilities for in-shore retrieval and compare them more rigorously with those of MODIS-A nFLH, OCI/OC3 and RGCI.

Finally, the critical test for the viability of KB HABs satellite retrieval techniques is the comparison of their ability to match retrieved values with concurrent in situ measurements. We therefore sought to extend the range of available in situ comparisons. This was done by seeking all available match-ups between VIIRS NN a_{ph_{443}} (and equivalent [Chla] retrievals) and in situ KB cell count measurements for the period 2012–2015. These comparisons showed that when the window between in situ observations and satellite measurements was reduced from daily, to 1 h and 0.5 h, the correlations greatly improved for the VIIRS NN and VIIRS RGCI retrievals against in situ match-ups. These showed respectable $R^2 = 0.82$ and $R^2 = 0.70$ respectively for the 0.5 h window. The comparisons also supported the choice of values for the F1 and F2 filter processes. Comparisons were also made comparing VIIRS NN retrievals identifying KB HABs, and in situ KB cell counts measured on the same date for 3 specific bloom events. All three confirmed the viability of the VIIRS NN retrieval approach, with no apparent false negatives or positives. Comparisons with other NN retrieval techniques are not so straightforward. Several of the previously reported NN techniques have used in situ measurements of IOPs, including sea surface temperatures and other physical parameters as inputs to NNs for bloom prediction, apparently with some success [57–59]. The more directly comparable NN technique reported, is the NN algorithm product for [Chla] retrievals in Case 2 waters [60–63] from the MERIS satellite, which is no longer operational. This appears to give comparable results, in European waters, to OCI/OC3 retrievals in open waters, and possibly higher accuracies in coastal waters [61]. This MERIS algorithm does not appear to have been tested for KB retrievals in the WFS. Tests of historical data for such retrievals may yield interesting results.

In conclusion, results to date with the VIIRS NN technique appear promising. However, more match-ups, including more detailed considerations of observation times and their connections to false positives or negatives, sample depths, distance to observation pixel centers and sub-pixel variability and impact of complex in-shore waters need to be studied. More comprehensive statistics need to be obtained on the overall efficacy of the approach and its comparisons to other techniques.

**Acknowledgments:** This work was partially supported by grants from NOAA through NOAA CREST, and NOAA JPSS. We would like to thank the NASA Ocean Color Processing Group for satellite imagery. We also thank NOAA HABSOS and Florida Fish and Wildlife Conservation Commission’s Fish and Wildlife Research Institute (FWRI) groups for making in situ KB data available to us.

**Author Contributions:** All authors contributed equally to this work.

**Conflicts of Interest:** The authors declare no conflict of interest.

**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
Appendix A

A1. NN Algorithm Background and Directions for Implementation for a\textsubscript{ph443} Retrieval

Note: The material in this Appendix is presented in a manner intended to help those readers less familiar with the use of Neural Networks. In the interest of completeness, some of the material presented here repeats parts of Section 2 above.

A.1.1. Background

In the work presented, we applied a synthetically trained NN algorithm previously developed and reported by us [12–15] to solve the inverse problem [19,20] of retrieving physical variables, including a\textsubscript{ph443} from, VIIRS observations of R\textsubscript{rs} 486, 551 and 671 nm in the WFS. A background summary is given here.

The essence of the approach is the use of a standard multiband NN inversion algorithm that takes VIIRS R\textsubscript{rs} measurements at the 486, 551 and 671 nm bands as inputs, and produces as output the related Inherent Optical Properties (IOPs), namely: absorption coefficients of phytoplankton (a\textsubscript{ph443}) dissolved organic matter (a\textsubscript{g}) and non-algal particulates (a\textsubscript{dm}) as well as the particulate backscatter coefficient, (b\textsubscript{bp}) all at 443 nm. (As discussed in Section 2 above, it is only with the NN output of a\textsubscript{ph443} that we are concerned with here).
A.1.2. Synthetic Dataset

For development of the algorithm, available measurement data from the NASA bio-Optical Marine Algorithm Data set [21], (NOMAD) was used as a basis for establishing a range of possible IOP parameters and examining potential relationships between them. The full range of parameters, applicable to oceanic and coastal waters, and the relationships established between them were used as inputs to a bio-optical model [14,22,23]. This model, Figure 1a, with four components: water, phytoplankton, colored dissolved organic matter (CDOM) and non-algal particulates (NAP) as inputs, was constructed to create a synthetic database consisting of 20,000 IOP combinations covering the entire observable range of IOP parameters. These IOPs were then used with a HydroLight based [36], parameterized forward model described in Lee et al., 2002 [23], to generate corresponding sets of Rs values at 486, 551 and 671 nm for VIIRS, and at 488, 555 and 667 nm for MODIS-A, associated with each of the 20,000 IOPs (and the related different combinations of the four components) [12–15]. The range and variability of IOPs used in our bio-optical modeling is well represented in the literature [22–35] and covers both case 1 and case 2 waters.

A.1.3. NN Training

50% (10,000) of these sets of Rs and related IOPs were then used [12–15], to train a VIIRS NN. Figure 1b, to model the relationship between input Rs values at the 486, 551 and 671 nm bands for VIIRS and the related IOPs $a_{ph}$, $a_{g}$, $a_{d}$, $bb_p$ coefficients, all at 443 nm, which is at the peak of phytoplankton absorption spectrum, and thus exhibits most variation in natural waters, (and where $a_g$ and $a_d$ are mutually constrained through an empirically derived relationship based on a study of NOMAD values) [12–15,21]. Figure 1b shows the VIIRS NN Architecture used. It is a one-hidden-layer multilayer perceptron (MLP) with six neurons at the hidden layer. Since this architecture was able to model our problems well, we did not further explore how this relationship might be improved by increasing the number of neurons or layers.

A back propagation learning procedure among a statistically representative synthetic dataset, based on NOMAD parameters for off shore and coastal waters was used to retrieve $a_{ph443}$ given the measured Rs values at three VIIRS bands 486, 551 and 671 nm. By suitable training, the back propagation learning procedure estimates the underlying complex relationships between the inputs and outputs by retrieving the NN weighting parameters and biases using standard minimization procedures on the residual output errors [64–69].

A.1.4. Testing the NN

The VIIRS NN algorithm was tested and evaluated using several statistical indicators. This was done initially on half the simulated data-set not used in the training, then on field data from the NOMAD. Finally, it was also tested on data from our own field campaign in the Chesapeake Bay during 2013, which represented well the typical range of water optical properties and chlorophyll-a concentrations in coastal regions. Shown below are the retrievals that were obtained against our in situ field measurements.

Figure A1,2 show the values retrieved using the VIIRS NN algorithm with in situ radiometer field measurements [14], of Rs at 486, 551 and 671 nm, plotted against the values of $a_{ph443}$ m$^{-1}$ (A1), and $bb_p(443)$ m$^{-1}$ (A2) measured by in situ WETLabs instrumentation [70]. These showed reasonably good correlations and demonstrated the potential efficacy of the VIIRS NN retrievals from in situ measured Rs values that are not impacted by atmospheric and surface effects. It provided impetus for testing NN applications for VIIRS HAB retrievals in the WFS, provided atmospheric corrections are adequate to give correct VIIRS derived Rs values at the 486, 551 and 671 nm bands.
A2. Retrieval of \( a_{\text{ph}443} \) from \( R_{\text{rs}486} \), \( 551 \) and \( 671 \) nm Using NN

This section describes the steps required for the NN algorithm implementation, to be used for retrieving \( a_{\text{ph}443} \) from \( R_{\text{rs}} \) 486, 551 and 671 nm. The related calculations can be performed using a MATLAB script.

Data transformation and standardization: It should be noted that inputs of the NN were normalized and transformed into the log domain and standardized [69] by subtracting the mean and dividing by the standard deviation before the training stage. Therefore outputs need to be denormalized in the retrieval stage to obtain meaningful results.

The following steps shows NN retrieval:

I Retrieval of NN \( a_{\text{ph}443} \) in (m\(^{-1}\)) units, is obtained by denormalizing the NN output to obtain Equation (a1). below:

\[
 a_{\text{ph}443} = 10^{(\sigma_{\text{o}} \hat{a}_{\text{ph}443} + \mu_{\text{o}})} \quad (a1)
\]

where \( \hat{a}_{\text{ph}443} \) is the normalized phytoplankton absorption at 443 nm obtained by NN and carries out the final weighted sum activity and biases associated with the NN output \( a_{\text{ph}443} \) m\(^{-1}\) and inputs \( R_{\text{rs}486}, R_{\text{rs}551} \) and \( R_{\text{rs}671} \) nm shown below in (Equation (a3)). The \( \mu_{\text{o}} (\lambda) \) and \( \sigma_{\text{o}} (\lambda) \) are the mean and standard deviation of the NN \( \hat{a}_{\text{ph}443} \) outputs respectively, obtained from the simulated datasets, and shown in Table A1.

II Neural Network parameters needed for the above calculations:

The first step shown in (Equation (a2)) below, requires input reflectance’s \( R_{\text{rs}} (\lambda) \), expressed as base 10 logarithm, standardized by removing, from each \( R_{\text{rs}} (\lambda) \), mean (\( \mu_{\text{i}} (\lambda) \)) of input values, of the simulated dataset, and then scaling the difference by the standard deviation (\( \sigma_{\text{i}} (\lambda) \)) of input values, of the simulated dataset, as shown in Equation (a2). Table A1 shows the relevant mean and standard deviation input (for each corresponding \( R_{\text{rs}} \) wavelength) and output values. (This procedure is also applied as the NN is trained using normalized \( R_{\text{rs}} (\lambda) \) values, so that it is equally sensitive to all inputs, avoiding conditioning problems).

<table>
<thead>
<tr>
<th>( \log_{10}(R_{\text{rs}486}) )</th>
<th>( \log_{10}(R_{\text{rs}551}) )</th>
<th>( \log_{10}(R_{\text{rs}671}) )</th>
<th>( \log_{10}(a_{\text{ph}443}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_{\text{i}} )</td>
<td>-2.2513</td>
<td>-2.4802</td>
<td>-3.4322</td>
</tr>
<tr>
<td>( \sigma_{\text{i}} )</td>
<td>0.1862</td>
<td>0.3456</td>
<td>0.5904</td>
</tr>
</tbody>
</table>
The variable \( t \), in the NN hyperbolic sigmoid activation function, \( \frac{2}{1+e^{-t}} - 1 \) is used to transfer the weighted sum activity of each computational element (neuron), which carries out the final weighted sum activity and biases associated with the trained NN input, is given by Equation (a2).

\[
t = \begin{bmatrix}
-0.0026 & 0.7735 & 0.1217 \\
0.6908 & -1.0168 & -0.3926 \\
0.2805 & 0.4950 & -1.7261 \\
-0.4861 & 1.3790 & -0.7815 \\
-0.2008 & 0.4675 & -0.0311 \\
-0.0940 & -0.0076 & 0.0165
\end{bmatrix} \ast \begin{bmatrix}
\log_{2}(R_{1}^{443}) - \log_{2}(R_{2}^{443}) \\
\log_{2}(R_{1}^{443}) - \log_{2}(R_{3}^{443}) \\
\log_{2}(R_{2}^{443}) - \log_{2}(R_{3}^{443}) \\
\end{bmatrix} + \begin{bmatrix}
2.2272 \\
-2.4660 \\
2.4989 \\
-0.5527 \\
-0.2028 \\
0.1321
\end{bmatrix} \tag{a2}
\]

\( \hat{a}_{ph443} \) is in turn related to \( t \), by Equation (a3).

\[
\wedge \hat{a}_{ph443} = \begin{bmatrix}
0.1410 & -0.6780 & -0.4435 & 0.0682 & 0.6546 & 0.3814 \\
-0.2470 & -0.8675 & -0.2453 & 0.0026 & 0.8058 & 0.6035 \\
0.1087 & 0.2762 & -0.0687 & -0.9550 & 2.2780 & 0.6473 \\
1.7204 & -0.0319 & -0.4128 & -0.1441 & 0.9679 & -1.2495
\end{bmatrix} \ast \left( \frac{2}{1+e^{-t}} - 1 \right) + \begin{bmatrix}
-0.2646 \\
-0.2638 \\
-0.1398 \\
-0.1019
\end{bmatrix} \tag{a3}
\]

This will result in a specific value of \( \hat{a}_{ph443} \) for a set of \( R_{s}486, R_{s}551 \) and \( R_{s}671 \) nm values. This value along with \( \mu_{o}(\lambda) \) and \( \sigma_{o}(\lambda) \) from Table A1 can now be used to complete evaluation of \( \hat{a}_{ph443} \) m\(^{-1}\) Equation (a1).

References


15. El-habashi, A.; Ahmed, S. 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. *Proc. SPIE* **2015**. [CrossRef]


19. Aires, F.; Prigent, C.; Rossow, W.B.; Rothstein, M. A 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. *J. Geophys. Res.* **2001**, *106*, 14887–14907. [CrossRef]


47. Wynne, T.T.; Stumpf, R.P.; Tomlinson, M.C.; Ransibrahmanakul, V.; Villareal, T.A. Detecting *Karenia brevis* blooms and algal resuspension in the western Gulf of Mexico with satellite ocean color imagery. *Harmful Algae* 2005, 4, 992–1003. [CrossRef]


49. Stumpf, R.P.; Culver, M.E.; Tester, P.A.; Tomlinson, M.; Kirkpatrick, G.J.; Pederson, B.A.; 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. [CrossRef]


60. Schiller, H.; Doerffer, R. Neural network for emulation of an inverse model operational derivation of Case II water properties from MERIS data. Int. J. Remote Sens. 1999, 20, 1735–1746. [CrossRef]