Satellite Ocean Color Based Harmful Algal Bloom Indicators for Aquaculture Decision Support in the Southern Benguela

Marié E. Smith * and Stewart Bernard 1,2

1 Earth Observation, CSIR, Cape Town, South Africa, 2 Department of Oceanography, University of Cape Town, Cape Town, South Africa

The aquaculture industry of southern Africa faces environmental threats from harmful algal blooms (HABs), which have the potential to cause devastating economic losses (Pitcher et al., 2019). Satellite earth observation offers a systematic and cost-effective method for operational monitoring of HABs over large areas. Whilst the chlorophyll-a concentration ([Chl-a]) product, often used as a proxy for phytoplankton biomass, can be used to indicate high biomass blooms (elevated biomass against a background signal of 5–10 mg Chl m⁻³), there is a clear need for value-added products that not only alert on bloom presence, but also on the bloom type and persistence. This study demonstrates the identification of different phytoplankton communities that can feasibly be identified in bloom concentrations from space, relevant to the aquaculture industry of South Africa. In terms of water-leaving reflectance, 76 % of the variance in the red and NIR spectral region is significantly positively correlated to phytoplankton abundance, [Chl-a], and the maximum line height (MLH) (defined as the height of the maximum reflectance peak above a baseline between 665 and 753 nm). The MLH is related to dominant phytoplankton types derived from phytoplankton count data, in order to identify thresholds which represent blooms that pose a high hypoxia and/or toxicity risk; whilst 0.0016 < MLH < 0.003 represent low to moderate concern mixed assemblage blooms, MLH > 0.003 has a strong likelihood of indicating high biomass dinoflagellate or Pseudo-nitzschia blooms. These techniques are routinely used by the aquaculture industry in South Africa for decision support and risk mitigation. The high biomass nature of the South African regional waters provide strong assemblage-related spectral variability, which can be exploited with the spectral bands of OLCI and MERIS. Application to these sensors not only ensures future monitoring capability, but also allows the creation of a historical risk climatology that can guide the site selection of industries sensitive to the presence of HABs, such as aquaculture facilities and desalination plants.

Keywords: ocean color, remote sensing, HABs, aquaculture, OLCI, MERIS, South Africa, Southern Benguela
1. INTRODUCTION

Aquaculture is a burgeoning industry in South Africa and plays a vital role in the country’s blue economy. The marine aquaculture sector centers around mussel *Mytilus galloprovincialis*, Pacific oyster *Crassostrea gigas* and abalone *Haliotis midae* farming (DAFF, 2017), with most facilities situated along the west coast in close proximity to the productive Benguela current upwelling system.

Most mussel and oyster farms utilize in-water culture methods such as rafts and cages (DAFF, 2017) located in Saldanha Bay (Figure 1). Abalone are farmed commercially along the west and southwest coasts of the country, with the majority of abalone operations situated near Walker Bay; the most common production methods utilize land-based flow-through systems (Urban-Econ Development Economists, 2018) which necessitates a close proximity to the ocean to allow large volumes of sea-water to be pumped up to the farm for optimal water exchange, temperature control, and removal of metabolic waste.

The global marine aquaculture sector faces environmental threats from harmful algal blooms (HABs), with impacts from these events amounting to approximately 8 $billion/yr (Brown et al., 2019). Within the South African aquaculture industry the HAB-related risk factors and mitigation strategies differs within the various sub-sectors. Whilst the herbivorous abalone are at risk of physical damage and paralysis attributed to some dinoflagellate species (e.g., Pitcher et al., 2019), the filter-feeders (i.e., mussels and oysters) are vulnerable to growth arrest (Pitcher and Calder, 2000) and the accumulation of biotoxins which affects their safety of consumption and can cause poisoning syndromes in humans. On a larger environmental scale, some non-toxic dinoflagellate blooms can result in marine mortalities and anoxia following the collapse of blooms with very high biomass (e.g., Ndhlouvu et al., 2017).

Routine management and risk assessment at aquaculture facilities includes monitoring the flesh of mussels and oysters for specific biotoxins and regular phytoplankton counts of water samples. Counts include total abundance counts and HAB species monitoring focusing on toxic dinoflagellates known to cause paralytic shellfish poisoning (PSP) (e.g., *Alexandrium* spp.) and diarrhetic shellfish poisoning (DSP) (e.g., *Dinophysis* spp.), diatoms known to cause amnesic shellfish poisoning (ASP) (e.g., some *Pseudo-nitzschia* spp.), as well as dinoflagellates known to produce yessotoxins (e.g., *Lingulodinium polyedrum*).

HABs have the potential to cause devastating economic losses in the aquaculture and fisheries industries. The Saldanha Bay mussel aquaculture industry was first affected in 1994 due to PSP (Pitcher et al., 1994), while the presence of brown tides in 1997–1999 resulted in reduced growth rates of the filter feeding bivalves and 80 % reduction in monthly sales (Probyn et al., 2001); in 2015 farms were closed 13 times due to the presence of biotoxins in shellfish flesh above acceptable regulatory limits (DAFF, 2017). Dinoflagellate blooms have previously impacted wild and farmed abalone (Pitcher et al., 2001; Botes et al., 2003) in South Africa, even leading to mortalities of wild adult abalone (Horstman et al., 1991); abalone have been known to contain paralytic shellfish toxins following some dinoflagellate blooms (Harwood et al., 2014; Hallegraeff and Bolch, 2016). HABs can also pose a threat to the physical condition of sardines and associated fisheries (Van der Lingen et al., 2016). The decay of high biomass dinoflagellate blooms have often lead to marine mortalities and mass rock-lobster strandings in the St Helena Bay region (e.g., Pitcher et al., 2011, 2014) with losses valued up to 50 million US dollars in some cases (Ndhlouvu et al., 2017).

Ocean color remote sensing provides a cost-effective and valuable tool in the detection and monitoring of various types of phytoplankton blooms (see references in Blondeau-Patissier et al., 2014). The most common method, using the concentration of Chlorophyll *a* ([Chl-*a*]) as a proxy for biomass, has often been used to define blooms as anomalous [Chl-*a*] above a pre-determined threshold (e.g., Stumpf et al., 2003). [Chl-*a*] has traditionally been used to detect phytoplankton blooms and HABs in the southern Benguela (Bernard et al., 2014; Smith and Bernard, 2018) where concentrations in the coastal waters are known to vary from <1 mg m$^{-3}$ in newly upwelled water (Barlow, 1982) to well over 100 mg m$^{-3}$ in bloom conditions (Pitcher and Weeks, 2006). Although [Chl-*a*] is routinely derived from satellite reflectance using regionally optimized algorithms (e.g., Smith and Bernard, 2018), it does not provide direct information about inherent phytoplankton-related risk.
The optical environment off the west coast of South Africa can be described as phytoplankton dominated, with other constituents (e.g., colored dissolved organic matter and suspended inorganic material) contributing relatively little to the water-leaving reflectance signal (Bernard et al., 2009). The blue-green water-leaving reflectance signal generally dominates in low biomass open ocean environments, which is considered to be \( [\text{Chl-a}] < 1 \text{ mg m}^{-3} \) in the context of this study. As the phytoplankton biomass increases, the reflectance signal is increasingly affected by a combination of the peaks of Chl-a absorption near 465 and 665 nm, the Chl-a fluorescence peak near 685 nm, as well as increased phytoplankton-related backscattering and the absorption of water; these effects result in the shift of the dominant reflectance signal to the red-NIR spectral region at Chlorophyll a concentrations ([Chl-a]) of approximately \( >15 \text{ mg m}^{-3} \) (Robertson Lain et al., 2014).

It should be noted that the term “bloom” and “HAB” should be considered within a specific ecological and environmental context; some HABs can be toxic at low biomass or do not manifest as high [Chl-a], and there are generally no set abundance values to define when a HAB species is considered to be a “bloom” (Gilbert et al., 2018). When contextualized within the highly productive waters of the southern Benguela, a phytoplankton “bloom” needs to be identifiable against a background biomass signal of approximately [Chl-a] of 5 to 10 mg m\(^{-3}\) (Demarcq et al., 2003); since a variety of different phytoplankton types commonly reach [Chl-a] of 20–50 mg m\(^{-3}\) and above (Bernard et al., 2014), a detection technique focusing on the red/NIR was deemed most appropriate for regional harmful bloom identification.

Spectral band difference algorithms are often used to relate the reflectance peak in the red/NIR to phytoplankton biomass, the most well-known version of which is arguably the fluorescence line height (FLH) (Letelier and Abbott, 1996; Gower et al., 1999); FLH provides a quantification of the height of the Chl-a fluorescence peak above a baseline formed by the Chl-a absorption trough near 665 and a NIR wavelength (usually near 750 nm). Several studies have used a variant of this line height detection method, whether on its own or in combination with other optical properties (e.g., backscattering) or true color imagery, for the detection of HABs in coastal waters using ocean color remote sensing (Gower et al., 2005; Ryan et al., 2008; Matthews et al., 2012; Al Shehhi et al., 2013; Ghanea et al., 2016).

This study aims to relate spectral features of water-leaving reflectance in the red-NIR directly to phytoplankton types of particular concern to the marine aquaculture industry of South Africa. We focus on application to reflectance data from the MEdium Resolution Imaging Spectrometer (MERIS) and the Ocean and Land Colour Imager (OLCI), as both sensors have spectral resolution in the red-NIR region and high (300 m) spatial resolution. The objective is to determine probabilistic ecosystem-contextualized identifiers for waters dominated by either dinoflagellates or the diatom *Pseudo-nitzschia* (PN) as these are high risk HAB types that offer distinct ocean color signals.

## 2. MATERIALS AND METHODS

*In situ* water samples were collected at a station in St Helena Bay approximately 4 km off of Lambert’s Bay (32.0845°S 18.2691°E) in late summer (between February and April) of 2004–2008. Chlorophyll a concentration was measured by fluorometric analysis (Holm-Hansen et al., 1965) using 90% acetone with the use of a Turner Designs 10-AU Fluorometer according to accepted protocols (Knap et al., 1996; Mueller et al., 2003). Phytoplankton samples were taken at the surface, fixed in buffered formalin to a concentration of 0.5%, and counted using the Utermöhl method (Hasle, 1978). Count data were grouped into diatoms, dinoflagellates, flagellates, ciliates, and coccolithophores; PN was treated separately, in an attempt to determine unique spectral characteristics. A >50 % abundance threshold was used as the primary simplistic phytoplankton population metric.

In-water radiometric measurements were made with a hyperspectral Tethered Atlanttic Radiometric Buoy (TSRB); further details on measurements, processing, and uncertainties can be found in Smith and Bernard (2018). The *in situ* radiometric data (\( N = 68 \)) were resampled to MERIS/OLCI wavelength bands centered at 665, 681.25, 708.75, and 753.75 nm.

A line height (or spectral band difference) algorithm, similar to the fluorescence line height (Gower et al., 1999), with a baseline formed by the water-leaving reflectance (\( \rho_w \)) between 665 and 753 nm was applied to all spectra. In the case of the hyperspectral *in situ* data the remote sensing reflectance (\( R_n \)) were converted to \( \rho_w \) by multiplying spectra by \( \pi \) (Antoine and Morel, 2005) prior to application of the line height algorithms. The line heights at both 681 (\( LH_{681} \)) and 709 nm (\( LH_{709} \)) were calculated as follows:

\[
LH(\lambda) = \rho_w(\lambda) - \rho_w(665) - [\rho_w(753) - \rho_w(665)] \times \left( \frac{\lambda - 665}{753 - 665} \right)
\]

The maximum line height (MLH) was calculated as follows:

\[
MLH = \max[LH_{681}, LH_{709}]
\]

The reflectance peak at 681 nm is generally associated with Chl-a fluorescence emission; however, at higher biomass this peak shifts to longer wavelengths due to the combined effects of increased phytoplankton absorption and backscattering, as well as pure water absorption. The ratio of \( LH_{709} \) to \( LH_{681} \), also known as the line height ratio (LHR) (Tao et al., 2011), provides an indication of this red shift, and was calculated as follows:

\[
LHR = \frac{LH_{709}}{LH_{681}}
\]

Principal component analysis (PCA) was applied as an exploratory data analysis step in order to assess the variance structure within the dataset. This analysis technique reduces the dimensionality of a dataset by breaking it down into a set of geometrically independent (orthogonal) modes of oscillation.
which represent all the variability in the data (Craig et al., 2012). PCA was performed on the resampled and standardized (i.e., removing the mean and scaling to unit variance) reflectance data using eigenvalue decomposition of the data covariance matrix. The scores of the first three principal components (modes) of variance were use in correlation analysis with the following variables: [Chl-a], MLH, LHR, the percentage compositions by abundance of diatoms, PN, dinoflagellates, flagellates, ciliates, coccolithophores, and the total cell counts for each sample. The authors note that although PN is a diatom it was assessed separately in an attempt to find an unique identification criteria given that is the only potentially toxic diatom genus appearing in the Benguela.

Satellite data from the Ocean and Land Colour Imager (OLCI) on board Sentinel-3A (processing baseline 2.23; IPF version 06.11) were obtained from the Copernicus online data access website (https://coda.eumetsat.int/) while data from the 3rd reprocessing for the Medium Resolution Imaging Spectrometer (MERIS) were obtained from the MERIS catalog and inventory (MERCI) website. In the case of both sensors the bright pixel (atmospheric) correction (Moore and Lavender, 2011) is universally applied over the coastal waters of the southern Benguela. The water-leaving reflectance ($\rho_w$) from the Level 2 data files were used in all calculations. The flags that were applied to maintain the quality of the data during the phytoplankton type detection algorithm application to MERIS data included “CLOUD,” “LAND,” uncorrected sun glint ("HIGLINT"), and reflectance confidence flags (“PCD1_13”); for OLCI data these included land and cloud flags (“LAND,” “CLOUD,” “CLOUD_AMBIGOUS,” “CLOUD_Margin”) missing, invalid or transmission errors (“INVALID,” “SUSPECT,” and “COSMETIC”), suspect atmospheric correction and saturated pixels (“AC_FAIL,” “SATURATED”), and unreliable sun glint correction flags (“RISKGLINT”).

3. RESULTS

The first three principal components accounted for 98.8 % of the total variance in the red-NIR region of the remote sensing reflectance ($R_{rs}$) dataset (Figure 2). The first mode, which accounts for 76% of the total variance within the dataset, represents an amplitude effect with a significant positive correlation to MLH, total cell count, [Chl-a] and LHR; this indicates that the biomass drives the magnitude of the $R_{rs}$ spectra in the red-NIR. The second mode indicates significant yet opposing spectral responses between $R_{rs}$ (665) and $R_{rs}$ (709) to variations in [Chl-a], LHR, MLH, and total cell counts. Mode two also had a significant positive correlation to the percentage coccolithophores in the sample; the highly scattering nature of these cells tend to increase the magnitude of the reflectance in the green, which in turn can partially mask some of the Chl-a absorption near 665 nm. The third mode, although contributing to a relatively small percentage of the total variance, is significantly negatively and positively related to the percentage composition of PN and dinoflagellates respectively; increases in the percentage of PN and dinoflagellates in water samples respectively are associated with increases in the LH681 and LH709, respectively, indicating a potential approach for the optical distinction of these two phytoplankton types.

Figure 3 shows the statistics of the MLH, LHR, and [Chl-a] associated with the dominance (i.e., >50 %) of diatoms, PN, and dinoflagellates respectively; only samples with total cell concentration over $10^6$ cells L$^{-1}$ were included in order to capture scenarios of likely phytoplankton bloom conditions. As the sample sizes were quite small, the Kruskal-Wallis H Test was used to compare the distributions of the three samples (diatom, PN, and dinoflagellates) for each variable (MLH, LHR, and [Chl-a]); this test found no significant differences between any of the samples. Bloom conditions dominated by diatoms tend to have a MLH < 0.0038; thus there is a high probability that conditions with MLH > 0.0038 is either PN or dinoflagellate dominated. For all three phytoplankton types approximately 75 % of the bloom samples had MLH > 0.0019; this was chosen as the lower threshold for mixed bloom conditions. Approximately half of the PN and 25 % of the diatom bloom samples had a MLH > 0.0027; this was chosen as the lower threshold for mixed bloom conditions that have a slightly higher potential for harm. All bloom conditions dominated by diatoms and PN displayed a LHR under 0.6; thus it is very likely that LHR > 0.6 would be dinoflagellate dominated. Similarly, valid samples of blooms dominated by either PN or other diatoms had maximum [Chl-a] under approximately 30 mg m$^{-3}$. Therefore, it is very likely that a bloom with [Chl-a] >30 mg m$^{-3}$ is dinoflagellate dominated; Bernard et al. (2014) also defined the lower end of the probabilistic range of dinoflagellate dominance as [Chl-a] above 30 mg m$^{-3}$. It should be noted that dinoflagellate dominance is entirely possible at lower [Chl-a] and/or LHR<0.6, but would most likely be associated with low total cell concentrations and related risk.

Some of the key values indicated above are used as baseline thresholds for a reflectance classification framework to determining phytoplankton types, which is presented in Table 2. An LHR > 0.6 is used as a probabilistic dinoflagellate identifier, while the MLH of 0.0019, 0.0027, and 0.0038 is used to represent increasing likelihood and potential severity of either dinoflagellates or PN blooms. Figure 4 shows the total cell counts and [Chl-a] that roughly corresponds to these in situ MLH thresholds; when using the regression lines of these figures there is a 63 % chance that a MLH of 0.0038 relates to total cell counts of approximately 6.4 million cells L$^{-1}$, and a 70% chance of it relating to [Chl-a] of approximately 23 mg m$^{-3}$. A small sample (N = 19) of coincident MERIS reflectance data were available to enable the comparison of satellite-derived MLH to in situ total cell counts and [Chl-a] (Figure 5); the coefficient of determination decreased from 0.63 to 0.52 for total cell counts, but increased from 0.70 to 0.74 for [Chl-a]. The thresholds for MLH were adjusted using the [Chl-a] associated with the original threshold values in Figure 3 and the new regression equation between satellite MLH and in situ [Chl-a]; the adjusted thresholds are shown in bold in Table 2, and are used for all satellite image classification.

Assessing the performance of algorithms designed to classify satellite data into discrete groups can be challenging, particularly
FIGURE 2 | Results from the principal component analysis of the in situ reflectance between 665 and 754 nm (N = 68). In the left panel the gray lines indicate individual \( R_{rs} \) samples, while the colored lines show the loadings of the top three modes of variability in the \( R_{rs} \) dataset. In the right panel the bars are color-coded according to the corresponding modes of variance to show the correlation with [Chl-a], maximum line height, line height ratio, percentage composition of diatoms, *Pseudo-nitzschia*, dinoflagellates, flagellates, ciliates, and coccolithophores, as well as the total cell counts per sample. The gray horizontal lines indicate the 95% significance level.

FIGURE 3 | Boxplots of the MLH, LHR, and [Chl-a] for samples with total cell concentration over 10^{6} cells L^{-1}. Each column represents the statistics associated with the dominance (i.e., >50% composition as relative abundance to the total cell count) of a given phytoplankton type, i.e., diatoms, *Pseudo-nitzschia*, and dinoflagellates. The horizontal lines of the boxes represent the 25th, 50th (median), and 75th percentiles, whereas the whiskers represent the valid minimum and maximum; outliers are indicated as diamonds. The colored dashed horizontal lines indicate the thresholds used in the initial probabilistic phytoplankton community classification.

when sample sizes are small, and often requires indicators other than the standard metrics (e.g., bias, RMSD) used for ocean color product validation (Melin et al., 2019); in these cases confusion/contingency/error matrices can be more useful (e.g., Carvalho et al., 2011; Wang and Hu, 2017). In the present study the performance of the satellite classification algorithm was assessed (in terms of correctly identifying pixels as either “bloom” or “non-bloom”) using Figure 5D as a reference; three confusion matrices (depicted in Table 1) were created to represent classification results of samples with in situ [Chl-a] above approximately 8.9, 14.2, and 22.1 mg m^{-3}, corresponding to MLH above the thresholds of 0.0016, 0.0022, and 0.003, respectively. Note that for the purpose of this assessment it is assumed that all samples above the specified [Chl-a] are true “blooms”. The classification accuracy is lowest (64%) at the low [Chl-a], with the satellite-derived classification indicating some false negatives; classification accuracy increases with increasing [Chl-a], with the highest accuracy obtained at [Chl-a] > 22.1 mg
m$^{-3}$. These results are based on very small sample sizes and are only used to provide an indication of the classification algorithm's operational limits.

Several studies have noted that the line height algorithms operating in the red/NIR can be affected by high concentrations of suspended sediment or atmospheric dust, producing apparent reflectance peaks in this spectral region (e.g., Zhao et al., 2015) or masking the reflectance peak signal (e.g., McKee et al., 2007; Gilerson et al., 2008). For the purposes of this study the area under the baseline of the MLH, i.e., the integral of the water-leaving reflectance values between the 665 and 753 nm wavebands, was used as a quality control measure for highly scattering (possibly inorganic) substances. The maximum MLH baseline integral of both the in situ and satellite-derived reflectance datasets was 0.5; thus for satellite application the classification algorithm was not applied to pixels where the integral was > 0.5 (i.e., these pixels are displayed as unclassified).

4. DISCUSSION

4.1. HABs in the Southern Benguela

The southern Benguela is a wind-driven, pulsed upwelling system forced by equatorward winds from the south Atlantic high pressure system, and modulated by low-pressure systems moving eastwards past the southern tip of Africa; these conditions supports elevated phytoplankton biomass over the wide continental shelf (Verheye et al., 2016) dominated by primarily large celled diatoms (Hutchings et al., 2012) that thrive in the nutrient-rich turbulent environment of upwelling systems (Sathyendranath et al., 2014). Succession within the system generally follows known conceptual frameworks (Margalef, 1987) where diatoms dominate during turbulent upwelling phases, followed by a shift to dinoflagellate dominance during quiescent periods. A decrease in upwelling-favorable winds toward the end of austral summer (between January and May) are usually associated with more frequent dinoflagellate-dominance in the near-shore waters of the southern Benguela during the latter stages of the upwelling season (Pitcher and Calder, 2000). HABs within the southern Benguela are largely attributable to dinoflagellates (Pitcher and Weeks, 2006). Although the prevalence of PN has been established in both the northern (Louw et al., 2016) and the southern (Fawcett et al., 2007) Benguela, there are no recorded impacts to the aquaculture industry (Pitcher et al., 2014).

Whilst HABs were considered to be relatively scarce along the southern coastline of South Africa prior to 1997 (Pitcher and Calder, 2000), several extensive dinoflagellate blooms, some consisting of previously unobserved species, have notably impacted the region in recent years (Pitcher et al., 2014); these events included *Gonyaulax polygramma* blooms that negatively affected physical condition of the regional sardine stock in 2011 (Van der Lingen et al., 2016), and blooms of *Lingulodinium polyedrum* impacting hundreds of kilometers along the south coast during 2013/2014 (Pitcher et al., 2014). Most notably for the aquaculture industry was the bloom co-dominated by *Lingulodinium polyedrum* and *Gonyaulax spinifera* at the end of 2016 which lead to the mortalities of over 250 tons of farmed abalone by February 2017 (Pitcher et al., 2019).
The occurrence and frequency of HABs are thought to be increasing worldwide, and within the context of a changing climate the global distribution and occurrence of different HAB species are likely to change in the future (Glibert and Burford, 2017). Pitcher et al. (2017) noted the continuously changeable nature of the species that constitute HABs in upwelling systems, and the inherently diverse threats posed to industries and humans relying on these systems. These concepts support the notion that the southern Benguela aquaculture industries requires adaptable and robust HAB monitoring strategies to safeguard the economic viability of these facilities.

4.2. Probabilistic Phytoplankton Community Classification (PPCC) Algorithm Functioning and Suitability

There are a multitude of methods to obtain information on phytoplankton functional types from remotely sensed ocean color data (see Sathyendranath et al., 2014), however many of these techniques were designed for oligotrophic and mesotrophic waters and/or operate in the blue-green spectral region. At lower biomass levels the spectral features in the blue-green wavelengths are potentially more useful for distinguish certain HABs from non-harmful blooms and other water types from an ocean color.
Examples 1, 2, and 3 represent samples with in situ [Chl-a] above 22.1, 14.2, and 8.9 mg m\(^{-3}\), respectively, as depicted in Figure 5D. For application to the MERIS and OLCI sensors, the traditional FLH utilizes wavebands centered at 665, 681, and 709 nm, while the Maximum Chlorophyll Index (MCI) uses 681, 709, and 753 nm (Gower et al., 2005). This relatively narrow positioning of the baseline and signal bands limits the application of the FLH to low-moderate biomass waters, whereas the MCI only functions in high biomass waters (i.e., [Chl-a] > 20 mg m\(^{-3}\) when the red-shift, associated with increasing phytoplankton biomass, produces a discernable reflectance peak in the red/NIR). Zhao et al. (2015) found that using a wide baseline modified FLH provided improved results compared to either the traditional FLH or MCI for qualitatively distinguish HABs from other blooms in the Arabian Gulf. The wide baseline and the dominant peak selection method are similar in functioning to the maximum peak-height (MPH) algorithm (Matthews et al., 2012) and the adaptive reflectance peak height (ARPH) algorithm (Ryan et al., 2014), which are both used in the operational quantification of different phytoplankton populations in eutrophic waters.

For application to the MERIS and OLCI sensors, the traditional FLH utilizes wavebands centered at 665, 681, and 709 nm, while the Maximum Chlorophyll Index (MCI) uses 681, 709, and 753 nm (Gower et al., 2005). This relatively narrow positioning of the baseline and signal bands limits the application of the FLH to low-moderate biomass waters, whereas the MCI only functions in high biomass waters (i.e., [Chl-a] > 20 mg m\(^{-3}\) when the red-shift, associated with increasing phytoplankton biomass, produces a discernable reflectance peak in the red/NIR). Zhao et al. (2015) found that using a wide baseline modified FLH provided improved results compared to either the traditional FLH or MCI for qualitatively distinguish HABs from other blooms in the Arabian Gulf. The wide baseline and the dominant peak selection method are similar in functioning to the maximum peak-height (MPH) algorithm (Matthews et al., 2012) and the adaptive reflectance peak height (ARPH) algorithm (Ryan et al., 2014), which are both used in the operational quantification of different phytoplankton populations in eutrophic waters.

4.3. Phytoplankton Community Identification Using Remote Sensing

Both MERIS and OLCI offer good spectral resolution in the red/NIR region as well as high spatial resolution (300 m) which is often necessary at the small spatial scale and near coastal aquaculture applications. With two satellites in orbit (Sentinel 3A and 3B), the OLCI sensor provides near daily coverage. In the generally eutrophic conditions of the southern Benguela it is useful to avoid the blue-green spectral region when using satellite-derived reflectance data, where the uncertainty resulting from aerosol extrapolation can be more extreme than in the red-NIR. The use of thresholds based on line height algorithms and

4.3. Phytoplankton Community Identification Using Remote Sensing

Both MERIS and OLCI offer good spectral resolution in the red/NIR region as well as high spatial resolution (300 m) which is often necessary at the small spatial scale and near coastal aquaculture applications. With two satellites in orbit (Sentinel 3A and 3B), the OLCI sensor provides near daily coverage. In the generally eutrophic conditions of the southern Benguela it is useful to avoid the blue-green spectral region when using satellite-derived reflectance data, where the uncertainty resulting from aerosol extrapolation can be more extreme than in the red-NIR. The use of thresholds based on line height algorithms and

| TABLE 1 | Confusion matrices describing the performance of the satellite classification model in terms of correctly identifying satellite-derived pixels as either “bloom” or “non-bloom.” |
|---|---|
| **Example 1** | **Satellite classification** |
| **In situ measurement** | **Non-bloom** | **Bloom** |
| **Non-bloom** | TN = 0 | FP = 0 |
| **Bloom** | FN = 0 | TP = 4 |
| **Classification accuracy:** | = 100% |

| **Example 2** | **Satellite classification** |
| **In situ measurement** | **Non-bloom** | **Bloom** |
| **Non-bloom** | TN = 0 | FP = 0 |
| **Bloom** | FN = 2 | TP = 6 |
| **Classification accuracy:** | = 75% |

| **Example 3** | **Satellite classification** |
| **In situ measurement** | **Non-bloom** | **Bloom** |
| **Non-bloom** | TN = 0 | FP = 0 |
| **Bloom** | FN = 5 | TP = 9 |
| **Classification accuracy:** | = 64% |

Examples 1, 2, and 3 represent samples with in situ [Chl-a] above 22.1, 14.2, and 8.9 mg m\(^{-3}\), respectively, as depicted in Figure 5D. For application to the MERIS and OLCI sensors, the traditional FLH utilizes wavebands centered at 665, 681, and 709 nm, while the Maximum Chlorophyll Index (MCI) uses 681, 709, and 753 nm (Gower et al., 2005). This relatively narrow positioning of the baseline and signal bands limits the application of the FLH to low-moderate biomass waters, whereas the MCI only functions in high biomass waters (i.e., [Chl-a] > 20 mg m\(^{-3}\) when the red-shift, associated with increasing phytoplankton biomass, produces a discernable reflectance peak in the red/NIR). Zhao et al. (2015) found that using a wide baseline modified FLH provided improved results compared to either the traditional FLH or MCI for qualitatively distinguish HABs from other blooms in the Arabian Gulf. The wide baseline and the dominant peak selection method are similar in functioning to the maximum peak-height (MPH) algorithm (Matthews et al., 2012) and the adaptive reflectance peak height (ARPH) algorithm (Ryan et al., 2014), which are both used in the operational quantification of different phytoplankton populations in eutrophic waters.

For application to the MERIS and OLCI sensors, the traditional FLH utilizes wavebands centered at 665, 681, and 709 nm, while the Maximum Chlorophyll Index (MCI) uses 681, 709, and 753 nm (Gower et al., 2005). This relatively narrow positioning of the baseline and signal bands limits the application of the FLH to low-moderate biomass waters, whereas the MCI only functions in high biomass waters (i.e., [Chl-a] > 20 mg m\(^{-3}\) when the red-shift, associated with increasing phytoplankton biomass, produces a discernable reflectance peak in the red/NIR). Zhao et al. (2015) found that using a wide baseline modified FLH provided improved results compared to either the traditional FLH or MCI for qualitatively distinguish HABs from other blooms in the Arabian Gulf. The wide baseline and the dominant peak selection method are similar in functioning to the maximum peak-height (MPH) algorithm (Matthews et al., 2012) and the adaptive reflectance peak height (ARPH) algorithm (Ryan et al., 2014), which are both used in the operational quantification of different phytoplankton populations in eutrophic waters.

4.3. Phytoplankton Community Identification Using Remote Sensing

Both MERIS and OLCI offer good spectral resolution in the red/NIR region as well as high spatial resolution (300 m) which is often necessary at the small spatial scale and near coastal aquaculture applications. With two satellites in orbit (Sentinel 3A and 3B), the OLCI sensor provides near daily coverage. In the generally eutrophic conditions of the southern Benguela it is useful to avoid the blue-green spectral region when using satellite-derived reflectance data, where the uncertainty resulting from aerosol extrapolation can be more extreme than in the red-NIR. The use of thresholds based on line height algorithms and

| TABLE 2 | The framework for probabilistic phytoplankton community classification (PPCC), based on thresholds of maximum line height (MLH) and line height ratio (LHR). |
|---|---|
| **MLH** | **LHR < 0.6** | **LHR > 0.6** |
| > 0.0016 (0.0019) | Mixed assemblage diatoms, low concern [pink] | Dinoflagellate dominated [green] |
| > 0.0022 (0.0027) | Mixed assemblage diatoms, moderate concern [blue] | Dinoflagellate dominated [green] |
| > 0.003 (0.0038) | High likelihood of Pseudo-nitzschia dominance [yellow] | High likelihood of dinoflagellate dominance, very high biomass [red] |

MLH thresholds based on in situ regression analysis are shown in brackets, and the adjusted values for application to satellite data are shown in bold text. The colors in square brackets represent the color-key used during PPCC algorithm application.
The probabilistic classification method was applied to four MERIS images (Figure 7) that coincided with the March 2006 field campaign where a *Pseudo-nitzschia* bloom was sampled off Lambert’s Bay (Fawcett et al., 2007). The PPCC correctly identified the presence of PN at the sampling station on the 12th of March, where [Chl-a] of 57.1 mg m⁻² and PN concentrations of 8 million cells L⁻¹ were measured *in situ*. Although the highest number of PN cells were measured on the 18th of March, the PPCC indicated only a high biomass mixed assemblage; this could potentially be due to the relatively lower *in situ* [Chl-a] (compared to the 12th of March), producing a poorer optical signal in the red/NIR. Both the 15th and 22nd of March coincided with decreased phytoplankton counts and [Chl-a], which were similarly reflected in the unclassified pixels over the sampling site. It is clear that the highest chance of successful classification is achieved under conditions of the highest biomass (i.e., on the 12th of March). Special precaution is also advised for regions classified as “high biomass mixed assemblages,” as they could likely contain high concentrations of PN.

### 4.4. Algorithm Limitations

This study is based on a relatively small *in situ* dataset (*N* = 68) with only a limited number of samples that included coincident radiometric measurements and phytoplankton counts; however, it did comprise a wide range of phytoplankton types and biomass concentrations enabling a first order determination of ecosystem-contextualized thresholds. Although these thresholds are based upon statistical indicators, the algorithm is not specifically meant to be a quantitative translation between MLH and cell counts or [Chl-a]. The aim was to provide simple intuitive map-based indication of probabilistic phytoplankton-related risk to the aquaculture industry of southern Africa. If/when more data become available these classification methods and detection accuracy could potentially be refined further.

The application of the PPCC to coincident satellite reflectance demonstrated that there was the highest likelihood of correct classification of phytoplankton community dominance at [Chl-a] > 23 mg m⁻³. Although it is possible that HAB identification could be more straight-forward at lower biomass levels under mono-specific bloom conditions, it is unlikely for any one species of phytoplankton to out-compete others under nutrient-rich upwelling conditions. DSP toxin producing species of the genus *Dinophysis*, usually *D. acuminata* or *D. fortii*, often form small components of blooms dominated by other dinoflagellates in the southern Benguela (Pitcher and Calder, 2000); these species also pose different threat levels to shellfish cultivation, as mussels are more susceptible to the accumulation of DSP toxins (Pitcher et al., 2011). Conversely it is also possible for...
some very high biomass dinoflagellate blooms to not result in any harmful impacts.

It should be noted that in situ and in vitro techniques far outweigh current remote sensing capabilities when it comes to phytoplankton identification at the species level. The strength of using remote sensing for HAB detection lies in the repeatability of measurements over the same location at higher spatial scales than is attainable by in situ methods. The utility of ocean color remote sensing for in HAB monitoring is most powerful when informed by coincident in situ information such as species-level phytoplankton identification and abundance, and toxicity. Remote sensing should ideally be utilized as part of a larger multi-scale monitoring approach: where a bloom has been identified as harmful, the PPCC method can aid in the continued monitoring of the bloom’s spatial extent, trajectory, and possible intensification or dissipation, thereby supporting decision making and risk mitigation processes at environmental and aquaculture management level.

The successful application of the algorithm to remote sensing data is dependent on the appropriate and successfully applied atmospheric correction and the resultant reflectance product quality. The algorithm functions in the red/NIR where the problems associated with aerosol correction are generally less than in the blue-green wavelengths. Line height algorithms are also generally less affected by absolute magnitude changes than ratios. Caution is advised when interpreting the Chl-a fluorescence signal from phytoplankton for quantitative determination of phytoplankton biomass, as the fluorescence efficiency of phytoplankton can be affected by various factors including taxonomy, physiology,
nutrient availability, light history, and temperature (Babin et al., 1996). Increased backscattering across the red/NIR, as might be caused by inorganic matter, could dampen the effect of the red/NIR reflectance peaks by increasing the reflectance of the baseline; studies have demonstrated that the FLH signal could be masked by non-algal materials in turbid waters (McKee et al., 2007; Gilerson et al., 2008).

4.5. Concluding Remarks and Future Outlook
This study represents a spectral classification scheme, applicable to both in situ and satellite reflectance, for the detection of phytoplankton communities relevant to the aquaculture industry of South Africa. Although the classification is primarily qualitative, it is based on species-related optical signatures and abundance data, and provides more direct risk-related information for aquaculture management than traditional maps of [Chl-a]. Future models could potentially incorporate environmental and/or nutrient information within the phytoplankton risk probability, as changes in these variables have been linked with bloom toxicity (e.g., Torres Palenzuela et al., 2019).

Whilst the classification system was contextualized for the southern Benguela, its utility is potentially appropriate to similar upwelling systems; for instance the northern Benguela is also known to experience frequent occurrences of toxic dinoflagellates (Dijerenge, 2015) and PN blooms (Louw et al., 2016), which could negatively impact regional marine aquaculture in Namibia. Following several years of severe drought in the western Cape province, there has been a rise in the number of planned desalination plants in the region; these facilities require appropriate phytoplankton monitoring practices, as both toxic and non-toxic algae can impact operations by clogging pre-treatment filters, causing saltwater reverse osmosis membrane fouling, and affecting the taste and odor of the water (Al Shehhi et al., 2017; Anderson et al., 2017). The PPCC could potentially be applied to 10 years of MERIS data and recent OLCI data to produce probability maps (e.g., Ryan et al., 2008) which could be used to guide future aquaculture and desalination site selection.

The routine high spatial information provided by the PPCC, used together with corroborative in situ phytoplankton cell counts, provides a powerful combination for operational HAB monitoring and daily decision support within the aquaculture industry. It is important that the limitations and strengths of the classifier be clearly delineated to users to ensure the appropriate level of response and mitigation, allowing different industries to use the information as they see fit.

DATA AVAILABILITY STATEMENT
The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS
MS and SB conceptualized the study. MS prepared figures, conducted analyses, and wrote the first draft of the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

ACKNOWLEDGMENTS
Grant Pitcher, Andre du Randt, and Trevor Probyn from the department of agriculture, forestry and fisheries for in situ data; GMES&Africa MarCOSouth and OCIMS for project funding. The original concept of this work was presented at the 2018 International Ocean Optics conference in Dubrovnik, Croatia (Smith and Bernard, 2018), and has been expanded upon in this manuscript.

REFERENCES
Al Shehhi, M. R., Gherboudj, I., and Ghedira, H. (2017). “Chapter 9 - Satellites-based monitoring of harmful algal blooms for sustainable desalination,” in Desalination Sustainability, ed H. A. Arafat (Elsevier), 341–366. doi: 10.1016/B978-0-12-809791-5.00009-2
Al Shehhi, M. R., Gherboudj, I., Zhao, J., Mezhoud, N., and Ghedira, H. (2013). “Evaluating the performance of MODIS FLH ocean color algorithm in detecting the Harmful Algae Blooms in the Arabian Gulf and the Gulf of Oman,” in Oceans-San Diego, 2013 (San Diego, CA: IEEE), 1–7.
Anderson, D. M., Boerlage, S. F. E., and Dixon, M. (2017). Harmful Algal Blooms (HABs) and Desalination: A Guide to Impacts, Monitoring, and Management. Paris: Intergovernmental Oceanographic Commission of UNESCO. doi: 10.25607/OBP-203
Antoine, D., and Morel, A. (2005). Atmospheric correction of the meris observations over ocean case 1 waters. MERIS ATBD 2.7.
Babin, M., Morel, A., and Gentili, B. (1996). Remote sensing of sea surface sun-induced chlorophyll fluorescence: consequences of natural variations in the optical characteristics of phytoplankton and the quantum yield of chlorophyll a fluorescence. Int. J. Remote Sens. 17, 2417–2448. doi: 10.1080/01431169608948781
Barlow, R. (1982). Phytoplankton ecology in the southern Benguela current. I. Biochemical composition. J. Exp. Mar. Biol. Ecol. 63, 209–227. doi: 10.1016/0022-0981(82)90179-4
Bernard, S., Pitcher, G., Evers-King, H., Robertson, L., Matthews, M., Rabagliati, A., et al. (2014). “Ocean colour remote sensing of harmful algal blooms in the Benguela system,” in Remote Sensing of the African Seas, eds V. Barale and M. Gade (Dordrecht: Springer), 185–203.
Bernard, S., Probyn, T., and Quirantes, A. (2009). Simulating the optical properties of phytoplankton cells using a two-layered spherical geometry. Biogeosci. Discuss. 6, 1497–1563. doi: 10.5194/bgd-6-1497-2009
Blondeau-Patissier, D., Gower, J. F., Dekker, A. G., Phinn, S. R., and Brando, V. E. (2014). A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans. Prog. Oceanogr. 123, 123–144. doi: 10.1016/j.pocean.2013.12.008
Botes, L., Sym, S. D., and Pitcher, G. C. (2003). Karenia crista sp. nov. and Karenia bicuneiformis sp. nov. (Gymnodiniales, Dinophyceae): two new Karenia species from the South African coast. Phycologia 42, 563–571. doi: 10.2216/i0031-8884-42-6-563.1
Brown, A. R., Lilley, M., Shutter, J., Lowe, C., Artioli, Y., Torres, R., et al. (2019). Assessing risks and mitigating impacts of harmful algal blooms on mariculture and marine fisheries. Rev. Aquacult. doi: 10.1111/raq.12403. [Epub ahead of print.]

Frontiers in Marine Science | www.frontiersin.org 11 February 2020 | Volume 7 | Article 61
Pitcher, G. C., Probyn, T. A., du Randt, A., Lucas, A., Bernard, S., Evers-King, H., et al. (2014). Dynamics of oxygen depletion in the nearshore of a coastal embayment of the southern Benguela upwelling system. *J. Geophys. Res. Oceans* 119, 2183–2200. doi: 10.1002/2013JC009443

Pitcher, G. C., and Weeks, S. J. (2006). The variability and potential for prediction of harmful algal blooms in the southern Benguela ecosystem. *Large Mar. Ecosyst.* 14, 125–146. doi: 10.1002/lre.20012

Probyn, T., Pitcher, G., Pienaar, R., and Nuzzi, R. (2001). Brown tides and mariculture in Saldanha Bay, South Africa. *Mar. Pollut. Bull.* 42, 405–408. doi: 10.1016/S0025-326X(00)00170-3

Robertson Lain, L., Bernard, S., and Evers-King, H. (2014). Biophysical modelling of phytoplankton communities from first principles using two-layered spheres: Equivalent Algal Populations (EAP) model. *Opt. Express* 22, 16745–16758. doi: 10.1364/OE.22.016745

Ryan, J., Davis, C., Tuñílarro, N., Kudela, R., and Gao, B.-C. (2014). Application of the hyperspectral imager for the coastal ocean to phytoplankton ecology studies in Monterey Bay, CA, USA. *Remote Sens.* 6, 1007–1025. doi: 10.3390/rs6021007

Ryan, J. P., Gower, J. F. R., King, S. A., Bissett, W. P., Fischer, A. M., Kudela, R. M., et al. (2008). A coastal ocean extreme bloom incubator. *Geophys. Res. Lett.* 35:L12602. doi: 10.1029/2008GL034081

Sathyendranath, S., Aiken, J., Alvain, S., Barlow, R., Bouman, H., Bracher, A., et al. (2014). *Phytoplankton Functional Types From Space*. Reports of the International Ocean Colour Coordinating Group (IOCCG); 15 (International Ocean-Colour Coordinating Group), 1–156.

Smith, M., and Bernard, S. (2018). Satellite ocean colour based harmful algal bloom identification for improved risk assessment and mitigation. *Environ. Sci.* 10.1002/essoar.10500045.1. [Epub ahead of print].

Stumpf, R., Culver, M., Tester, P., Tomlinson, M., Kirkpatrick, G., Pederson, B., et al. (2003). Monitoring *Karenia brevis* blooms in the Gulf of Mexico using satellite ocean color imagery and other data. *Harmful Algae* 2, 147–160. doi: 10.1016/S1568-9883(02)00083-5

Tao, B., Mao, Z., Lei, H., Pan, D., Shen, Y., Bai, Y., et al. (2015). A novel method for discriminating *Proorocentrum donghaiense* from diatom blooms in the East China Sea using MODIS measurements. *Remote Sens. Environ.* 158, 267–280. doi: 10.1016/j.rse.2014.11.004

Tao, B., Mao, Z., Wang, D., Lu, J., and Huang, H. (2011). “Remote sensing of the ocean, sea ice, coastal waters, and large water regions 2011,” in *Proceedings of SPIE*, Vol. 8175 (Prague).

Torres Palenzuela, J. M., González Vilas, L., Bellas, F. M., Garet, E., González-Fernández, Á., and Spyrakos, E. (2019). *Pseudo-nitzschia* blooms in a coastal upwelling system: Remote sensing detection, toxicity and environmental variables. *Water* 11:1954. doi: 10.3390/w11091954

Urban-Econ Development Economists (2018). *Abalone Feasibility Study. Final report*. Report prepared for the Department of Agriculture, Forestry and Fisheries, South Africa.

Van der Lingen, C., Hutchings, L., Lamont, T., and Pitcher, G. (2016). Climate change, dinoflagellate blooms and sardine in the southern Benguela Current Large Marine Ecosystem. *Environ. Dev.* 17, 230–243. doi: 10.1016/j.envdev.2015.09.004

Verbeeye, H. M., Lamont, T., Hoggett, J. A., Kreiner, A., and Hampton, I. (2016). Plankton productivity of the Benguela current large marine ecosystem (BCLME). *Environ. Dev.* 17, 75–92. doi: 10.1016/j.envdev.2015.07.011

Wang, M., and Hu, C. (2017). Predicting Sargassum blooms in the Caribbean Sea from MODIS observations. *Geophys. Res. Lett.* 44, 3265–3273. doi: 10.1002/2017GL072932

Zhao, J., Temimi, M., and Ghedira, H. (2015). Characterization of harmful algal blooms (habs) in the Arabian Gulf and the Sea of Oman using MERIS fluorescence data. *ISPRS J. Photogramm. Remote Sens.* 101, 125–136. doi: 10.1016/j.isprsjprs.2014.12.010

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Smith and Bernard. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.