How optically diverse is the coastal ocean?

F. Mélin, Vincent Vantrepotte

Remote Sensing of Environment 160 (2015) 235–251

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

How optically diverse is the coastal ocean?

F. Mélin [a,*], V. Vantrepotte [b,c]

[a] European Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability (IES), TP270, via Fermi 2749, Ispra 21027, Italy
[b] INSU-CNRS, UMR 8187, LOG, Laboratoire d'Océanologie et des Géosciences, Université Lille Nord de France, ULCO, 32 avenue Foch, 62930 Wimereux, France
[c] CNRS Guyane, USR 3456, 2 avenue G. Charlery, 97300 Cayenne, Guyane, France

ARTICLE INFO

Article history:
Received 6 June 2014
Received in revised form 23 January 2015
Accepted 26 January 2015
Available online 18 February 2015

Keywords:
Optical properties
Coastal regions
Classification
SeaWiFS
Ocean color

ABSTRACT

Coastal regions are a resource for societies while being under severe pressure from a variety of factors. They also show a large diversity of optical characteristics, and the potential to optically classify these waters and distinguish similarities between regions is a fruitful application for satellite ocean color. Recognizing the specificities and complexity of coastal waters in terms of optical properties, a training data set is assembled for coastal regions and marginal seas using full resolution SeaWiFS global remote sensing reflectance R_{RS} data that maximize the geographic coverage and seasonal sampling of the domain. An unsupervised clustering technique is operated on the training data set to derive a set of 16 classes that cover conditions from very turbid to oligotrophic. When applied to a global seven-year SeaWiFS data set, this set of optical water types allows an efficient classification of coastal regions, marginal seas and large inland water bodies. Classes associated with more turbid conditions show relative dominance close to shore and in the mid-latitudes. A geographic partition of the global coastal ocean serves to distinguish general optical similarities between regions. The local optical variability is quantified by the number of classes selected as dominant across the period, averaging 5.2 classes if the cases accounting for 90% of the data days are considered. Optical diversity is more specifically analyzed with a Shannon index computed with the class memberships. Regions with low optical diversity are the most turbid waters as well as closed seas and inland water bodies. Oligotrophic waters also show a relatively low diversity, while intermediate regions between coastal domain and open ocean are associated with the highest diversity, which has interesting connections with ecological features.

1. Introduction

As the interface between the oceans and the terrestrial ecosystems supporting the human population, coastal regions need to be thoroughly studied and understood. They harbor a large share of Earth's population with high densities (Small & Nicholls, 2003) and represent a resource for societies, but consequently they are under increasing pressure from anthropogenic origin. Increasing population in large cities and economic development, waste water discharge and other localized pollution, and habitat disruptions have already resulted in a degradation of coastal ecosystems (Halpern et al., 2008; Lotze et al., 2006). Human activities directly impact the food web structure and biodiversity of coastal regions through intense fishing (Stewart et al., 2010) and by favoring the invasion of alien species (e.g., Katsanevakis, Zenetos, Belchor, & Cardoso, 2013; Molnar, Gamboa, Revenga, & Spalding, 2008). The extension of the network of river impoundments modifies both the flow of fresh water and the amount of sediments reaching estuaries (Vörösmarty et al., 2003). Anthropogenic nutrient inputs to coastal zones also have a strong chemical signature (Galloway et al., 2008) leading to eutrophication and hypoxia phenomena (Diaz & Rosenberg, 2008; Voss et al., 2011). These effects can be compounded by the increase in greenhouse gas concentrations and its associated warming and acidification. Consequences of climate change for plankton species distribution or for the functioning of upwelling ecosystems have already been suggested (e.g., Bakun & Weeks, 2004; Beaugrand, Reid, Ibañez, Lindley, & Edwards, 2002).

In that context, coastal ecosystems and marginal seas need to be properly monitored to allow an improved understanding of their dynamics and the detection of changes in their properties, and to follow the impact of policies aiming at environmental protection. But while a global observing network is required, the actual sampling is very unevenly distributed in space or time: a large part of the coastal regions have been poorly sampled by optically relevant observations and some seasons are relatively ignored because of operational constraints. Remote sensing of ocean color has a role to play as a cost-effective tool for global and frequent observations that can be interpreted in terms of surface concentrations of chlorophyll-a (Chl*a*), suspended material or chromophoric dissolved organic matter (CDOM). However this global capability is to some extent questioned by the uneven

* Corresponding author.

http://dx.doi.org/10.1016/j.rse.2015.01.023

distribution of field data that are at the basis of empirical algorithms or that are used for the definition of parameters in semi-analytical bio-optical algorithms, like the Chl*a*-specific absorption coefficient or the exponential slope of CDOM absorption. For the same reason, validation results have a limited coverage particularly for remote sensing reflectance or inherent optical properties. This raises the question of the applicability of most algorithms as well as the validation statistics. This issue can be circumvented by the introduction of techniques of optical classification that use the spectrum of remote sensing reflectance R_{RS} to quantify how two water bodies are similar from the optical point of view (e.g., D'Alimonte, Mélin, Zibordi, & Berthon, 2003; Martin Traykovski & Sosik, 2003; Moore, Campbell, & Feng, 2001). With the proviso that algorithm definition or validation results obtained for one optical water type (or class) are applicable to any water body associated with that type, optical classification techniques represent a powerful avenue for an optimal and truly homogeneous use of ocean color remote sensing at global scale. This potential has already been exploited for regional or global applications (Lubac & Loisel, 2007; Mélin et al., 2011; Moore, Campbell, & Dowell, 2009; Moore, Dowell, Bradt, & Ruiz Verdu, 2014; Vantrepotte, Loisel, Dessailly, & Mériaux, 2012) by making use of in-situ data bases to define the optical water types of reference. This presents various advantages as field data are usually considered as having uncertainties lower than those of satellite data, and as the collection of field R_{RS} data is often accompanied by other measurements needed for the envisioned application (e.g., Chl*a* for the definition of an associated empirical algorithm). However in that case the optical variability covered by the classification is restricted to the range of the field data, which is a limitation for the creation of a set of optical water types that would evenly represent all regions and seasons.

The objective of this study is to use satellite data available for coastal/shelf waters and marginal seas to derive a set of optical water types encompassing the full extent of the optical variability found in these regions, information as yet not available. The focus on coastal regions is justified by their importance; often termed optically complex waters, they also contain a large share of the optical variability of natural waters in contrast to the fairly constrained variability found in most open-ocean waters (Morel & Maritorena, 2001). However, as will be seen below, optical water types typical of oligotrophic to mesotrophic waters are covered in the analysis. The main application of the study is to document the optical variability observed at global scale, to expose the optical similarity between regions, or to assess the degree of optical stability in a given region. The domain of study and the creation of the training data set are first presented. Then clustering and classification approaches are described, and finally the distribution of optical water types is documented. Finally, the optical diversity is quantified.

2. Data & methods

2.1. Satellite data and domain of study

All the Sea-viewing Wide Field-of-View Sensor (SeaWiFS, McClain et al. 1998) Level-1A data have been collected from the Ocean Biology Processing Group (OBPG) of the National Aeronautics and Space Administration (NASA) and processed with the SeaWiFS Data Analysis Software (SeaDAS, version 6, Fu, Baith, & McClain, 1998). This imagery is the so-called MLAC (Merged Local Area Coverage) acquisition with a resolution of ~1.1 km at nadir. Products were mapped onto a global domain with a sinusoidal projection and a spatial resolution of 1/48th-degree (approximately 2.3 km), and subsequent analyses were made with daily data. This spatial and temporal sampling is well adapted to capture the optical variability found in coastal waters, whereas a higher level of averaging would tend to smooth out peculiar spectral characteristics. A drawback of using the SeaWiFS MLAC data is that data acquisition has been uneven in space and time following the operation of receiving ground stations.

The focus of this work is on coastal regions and marginal seas. To isolate the part of the global ocean that responds to that vague definition, a set of arbitrary criteria was applied to define the domain of study. First, grid points were excluded if the shortest distance to the coast was larger than 200 km or if the bottom depth was deeper than 4000 m (9000 m for the region along the western coasts of South and Central America). Bathymetry is defined according to the General Bathymetric Chart of the Oceans (GEBCO) 1-minute gridded data set. Finally, some marginal seas, parts of which were excluded by these criteria, were restored to their full extent in the domain of analysis, e.g., the Indonesian Archipelago, the Chinese and Japan Seas, the Sea of Okhotsk, the Mediterranean Sea, the North Sea, the Gulf of Mexico, and the Hudson Bay. Very large lakes were also included (e.g., Great American, European and African Lakes, Lake Baikal, Caspian Sea) but the associated data are not included in the training process. The final coastal domain amounts to 12% of the Earth surface (or 17% of the ocean domain).

To retrieve regional statistics, the domain was split into distinct regions representative of marginal seas or known partitions of the coastal ocean (Fig. 1). In particular, this regional distribution was partly inspired by the biogeographic provinces of Longhurst (2006) and the Large Marine Ecosystems (LMEs) partition (Sherman & Hempel, 2009). Tables 1 to 7 provide the list of acronyms used to designate each region. The selected name does not necessarily reflect the entire geographic domain usually associated with the region but only the coastal/shelf part considered for the analysis.

The SeaWiFS data were processed for the interval 1998–2004 (7 years) which was a period of unrestricted distribution of LAC data by NASA for research purposes. Fig. 2 illustrates the total number of days with valid data that went into the analysis. The average number is 283 days over the 7-year period (standard deviation, s.d., 221 days) for the entire domain. This relatively small amount of days is explained by the use of MLAC imagery as explained above but is sufficient to conduct a global study. This coverage appears highly variable, having a maximum of 767 days for the Mediterranean Sea (MEDI) and a minimum of 6 days for the Laptev Sea (LAPT). The regions with the lowest coverage are found in the high latitudes or associated with frequent cloud or dust cover, such as the Gulf of Guinea (GUIN, average of 41 days). Some inland water bodies show a fairly low coverage (down to 12 days for Lake Baikal) but this is not a general feature (the American Great Lakes, GREL, count an average of 358 days of valid data). An additional element modulating the data availability is linked to the operations of the receiving ground stations.

2.2. Training data set

A training data set is needed to define a set of optical classes. To be manageable, this data set can only be a subset of the overall satellite data archive, yet it should be well representative of the optical variability found in natural coastal waters. The approach followed here aimed at maximizing the geographic coverage and seasonal sampling of the training data set. For each grid point within the domain, a list of days with valid R_{RS} spectra was built. Out of that list, five days were retained optimizing their dispersion along the calendar year. If a grid point had less than five days with valid data, they were all included in the training data set. This approach ensured that a location or a season with very few data were still represented in the training data set, or conversely avoided that regions with many valid data or that seasons with the most favorable atmospheric conditions dominated the training process. The final training data set amounts to 51 million spectra. The full extent of the oceans is not included in the creation of the training data set because the mesotrophic to oligotrophic waters would account for a large share of the data and weaken the ability of the clustering step to capture subtle optical variations in coastal regions.

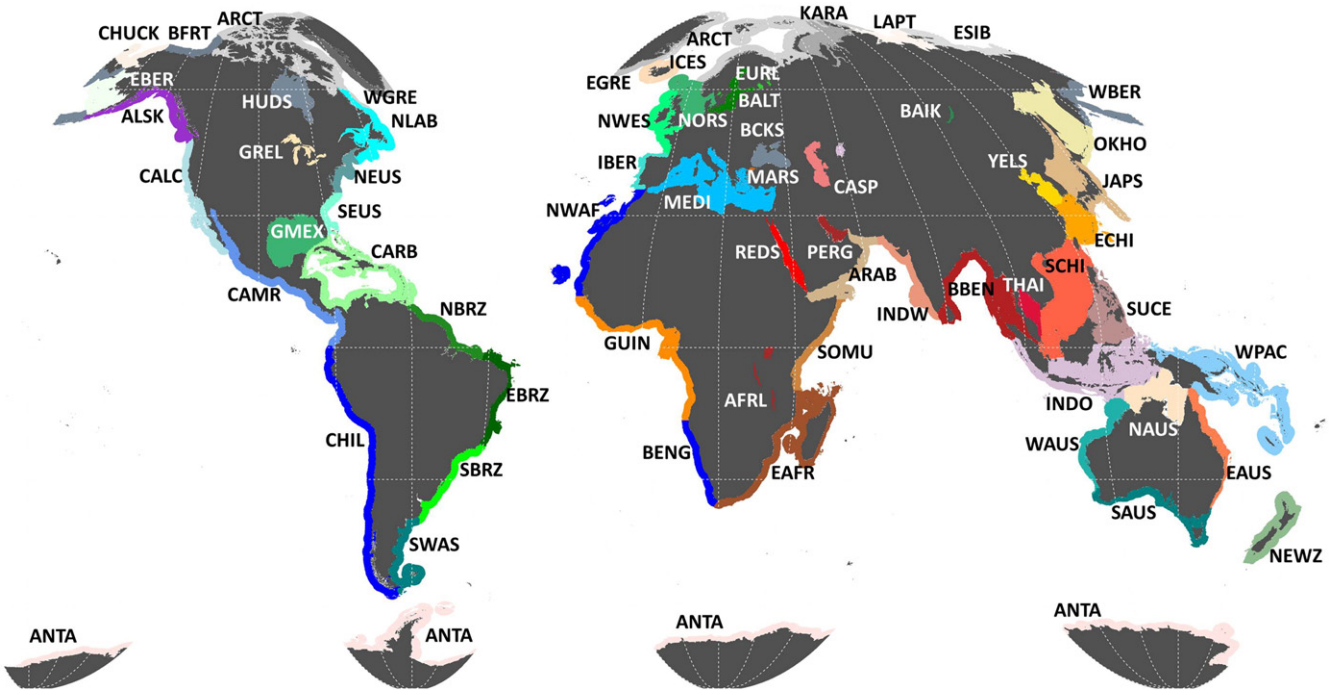

Fig. 1. Partition of the domain of analysis into regions, the acronyms of which are defined in Tables 1 to 7.

2.3. Clustering technique

In order to reduce the reflectance first order variability and focus on the reflectance spectral shape each R_{RS} spectrum was normalized prior to the data clustering by its integrated value (i.e., the surface below the spectrum, Lubac & Loisel, 2007; Vantrepotte et al., 2012) following the formula (for each wavelength λ):

$$r_n(\lambda) = \frac{R_{RS}(\lambda)}{\int_{\lambda_1}^{\lambda_2} R_{RS}(\lambda) d\lambda} \quad (1)$$

where r_n (in units of nm^{-1}) indicates the normalized spectrum obtained by integration between λ_1 (412 nm) and λ_2 (670 nm). This formula was computed by trapezoidal integration.

Normalized reflectance spectra composing the training data set were then analyzed using the ISODATA (Iterative Self-Organizing Data Analysis Technique, Jensen, 1996; Memarsadeghi, Mount, Netanyahu, & Le Moigne, 2007; ISODATA, 2014) clustering method. This unsupervised classification algorithm corresponds to a modified version of the *k*-means clustering. Different from the *k*-means procedure, ISODATA allows the number of clusters to be adjusted automatically during the iterative process which aims at merging similar clusters and splitting distant ones based on the spectral standard deviation and distance criteria. At the first iteration, cluster centers are randomly placed and spectra are assigned to the different clusters based on the shortest distance to cluster centers. At any subsequent iteration, clusters are split if the standard deviation is greater than a user-defined threshold (0.001), and are merged if the distance between them is lower than a user-defined threshold (0.00001). At each step of the procedure cluster distance and standard deviation are computed and iterations end when the number of spectra in each cluster changes by less than a defined threshold (5%). The minimum number of spectra composing a defined cluster was set to 1000 in order to avoid the consideration of unrealistic spectral shapes, in practice assigning outliers data to an unclassified category. Other parameters were kept at their default value (ISODATA, 2014). This method which presents a clear heuristic character has been shown to be well adapted to treat remote sensing data sets especially for land cover applications (Castillo Atoche, Carrasco Alvarez, Ortegón Aguilar, & Vázquez Castillo, 2013; McCullough, Loftin, & Sader, 2012; Walsh et al., 2008). The process returns the number of classes and

Table 1
Definition of regions' acronyms for high latitude regions. The dominant class (class with the most frequent occurrence of maximum membership) is indicated for each region, as well as the average Shannon index (see Eq. 6).

Acronym	Description	Dominant class	Shannon index
ARCT	Arctic Ocean	6	2.00
KARA	Kara Sea	1	1.36
LAPT	Laptev Sea	1	0.98
ESIB	East Siberian Sea	1	1.11
WBER	West Bering Sea	6	2.15
EBER	East Bering Sea	2	1.76
CHUK	Chukchi Sea	2	1.88
BFRT	Beaufort Sea	1	1.62
HUDS	Hudson Bay	7	1.68
WGRE	West Greenland Region	7	2.19
EGRE	East Greenland Region	8	2.19
ANTA	Antarctic Region	11	1.82

Table 2
As Table 1 for European seas and western Atlantic regions.

Acronym	Description	Dominant class	Shannon index
ICES	Iceland Shelf	6	2.26
BALT	Baltic Sea	1	1.11
NORS	North Sea	4	1.95
NWES	Northwest European Shelf	4	2.03
MEDI	Mediterranean Sea	14	1.92
MARS	Marmara Sea	3	1.69
BCKS	Black Sea	3	1.65
IBER	Iberian Upwelling	10	2.23
NWAF	Northwest African Upwelling	14	2.08
GUIN	Gulf of Guinea	7	2.16
BENG	Benguela Upwelling	4	1.98

Table 3
As Table 1 for eastern Atlantic regions.

Acronym	Description	Dominant class	Shannon index
NLAB	Newfoundland-Labrador	7	2.05
NEUS	Northeast US Shelf	4	2.03
SEUS	Southeast US Shelf	15	1.97
GMEX	Gulf of Mexico	14	1.99
CARB	Caribbean Sea	15	1.72
NBRZ	North Brazil Region	1	1.81
EBRZ	East Brazil Region	16	1.53
SBRZ	South Brazil Region	1	2.04
SWAS	Southwest Atlantic Shelf	4	1.98

the distribution of the reflectance spectra that are members of each class.

2.4. Classification technique

The classification consisted in calculating the class membership of an input spectrum for each class defined statistically by its average spectrum, μ, and covariance matrix, $\sum$, from the population of members that were associated with that class in the clustering process (where statistics μ and $\sum$ are calculated with log-transformed r_n; D'Alimonte et al., 2003; Mélin et al., 2011; Vantrepotte et al., 2012). Before classification, the reflectance spectrum R_{RS} was normalized by the area below the spectrum consistently with the training phase to produce r_n (Eq. 1), and was then log-transformed. The distance of the resulting input x to a given class ic was quantified by the squared Mahalanobis distance ${\Delta_{ic}}^2$:

$$ {\Delta_{ic}}^2(x) = (x-\mu_{ic})^T {\Sigma_{ic}}^{-1} (x-\mu_{ic}) \quad (2) $$

where T indicates the matrix transpose. The class membership p for class ic and input x was quantified as:

$$ p(ic,x) = 1-P\left({\Delta_{ic}}^2(x), n\right) \quad (3) $$

where P is the cumulative χ^2 distribution function with n degrees of freedom (Moore et al., 2001), in that case the number of wavelengths used for the classification (the 6 SeaWiFS bands). In practice, P was defined as:

$$ P\left({\Delta_{ic}}^2(x), n\right) = \frac{\gamma\left(n/2, {\Delta_{ic}}^2(x)/2\right)}{\Gamma(n/2)} \quad (4) $$

with Γ the Gamma function, and γ the lower incomplete Gamma function (Gallego, Cuevas, Mohedano, & García, 2013). The classification operates on log-transformed r_n in order to be more robust to outliers and in agreement with the log-normal hypothesis associated with natural waters (Campbell, 1995). A test supporting the use of log-normal statistics in the classification was made by computing the Mahalanobis distance between the log-transformed r_n members of each class within the training data set and the respective class center (auto-classification). For each class ic, the proportion of members associated with a Mahalanobis distance ${\Delta_{ic}}^2$ lower than 10.6 with respect to the class center is remarkably close to 90% (mean of 90.3% ± 0.9%, minimum 88.6%, maximum 91.3%), which is the percentage expected for a multi-normal distribution of size 6 (the number of bands) (Press, Teukolsky, Vetterling, & Flannery, 1992).

Table 4
As Table 1 for eastern Pacific regions.

Acronym	Description	Dominant class	Shannon index
ALSK	Gulf of Alaska	4	2.06
CALC	California Coastal Current	7	2.18
CAMR	Central America Coastal Province	11	2.29
CHIL	Peruvian/Chilean/Humboldt Current	8	2.10

Table 5
As Table 1 for western Pacific regions.

Acronym	Description	Dominant class	Shannon index
OKHO	Sea of Okhotsk	6	2.12
JAPS	Japan Sea	7	2.27
YELS	Yellow Sea	1	1.37
ECHI	Eastern China Sea	16	1.99
SCHI	Southern China Sea	15	1.92
THAI	Gulf of Thailand	10	1.91
INDO	Indonesian Seas	13	2.15
WPAC	Western Pacific	15	1.65
NAUS	Northern Australia	7	2.02
EAUS	Eastern Australia	15	1.87
NEWZ	New Zealand	11	2.16

2.5. Statistics

Once the class membership p was calculated for each day and bin, the normalized membership p^* was expressed for class ic as:

$$ p^*(ic) = \frac{p(ic)}{\sum_{ic=1}^{NC} p(ic).} \quad (5) $$

The classification scheme does not impose a sum of memberships equal to 1. For instance, a spectrum associated with optical conditions not well represented in the training data set could have low class membership for all classes. It is therefore important to illustrate the distribution of total memberships to assess the degree to which the classification scheme appropriately covers the optical variability of the considered water body. For each bin, the number of days when different thresholds of total memberships were reached was also recorded. Normalized memberships were used when comparing and combining data from different days or locations.

To distinguish the dominant class, the class of maximum membership was recorded for each bin and day, and these occurrences were accumulated in time over the period 1998–2004. For each bin, the class most frequently selected as the class of maximum membership, hereafter termed dominant class, was then determined. It should be mentioned here that this class may not always be associated with the highest average class membership (where the average is computed for each class over the period 1998–2004). Subsequently the occurrences of dominant class (number of days when a class was selected as the class of maximum membership) were averaged for each region to determine the dominant class for that region, reported in Tables 1–7. This is a general indicator that should be considered with caution since there is often a large spatial variability in the

Table 6
As Table 1 for Indian Ocean regions.

Acronym	Description	Dominant class	Shannon index
EAFR	Eastern Africa	14	2.04
SOMU	Somalia Upwelling	11	2.33
REDS	Red Sea	14	2.03
ARAB	Arabian Sea	6	2.35
PERG	Persian Gulf	2	1.56
INDW	Western India	2	1.94
BBEN	Bay of Bengal	12	1.96
WAUS	Western Australia	12	2.06
SAUS	Southern Australia	10	2.08

Table 7
As Table 1 for inland water bodies.

Acronym	Description	Dominant class	Shannon index
GREL	American Great Lakes	3	1.50
EURL	European Lakes	1	0.73
CASP	Caspian Sea	2	1.35
EAFR	African Lakes	3	1.49
BAIK	Lake Baikal	2	1.29

dominant class for many regions. Furthermore, the class of maximum membership should not be construed as exclusive: other classes might have a membership value close to the maximum.

For each bin, the number of classes that were selected at least once as the class of maximum membership was recorded too. As this indicator can be affected by outliers or rare conditions, the number of classes that were selected as the class of maximum membership while accounting for 90% of the total number of days was also recorded for each bin.

3. Results

3.1. Definition of the optical classes

The clustering analysis applied on the training data set led to the definition of 16 classes which average normalized spectra (r_n) are represented in Figs. 3 and 4 (statistics defining the classes are available upon request). Fig. 4 also shows the average raw (i.e., without normalization) spectra R_{RS}. Note that only 0.01% of the 51 million spectra composing the training data set were registered as unclassified; examining these cases often revealed erratic spectral shapes that might result from difficulty in the atmospheric correction process.

The average r_n spectrum for class 1 is characterized by a strong signal in the green and red parts of the spectrum (with a maximum at 555 nm) and low r_n at the blue wavelengths (Fig. 3a). This spectral shape is typical of very turbid water masses associated with a large amount of mineral particles and dissolved organic matter (Babin & Stramski, 2004; Vantrepotte et al., 2012). Spectral shapes associated with classes 2 to 7 are also representative of r_n spectra usually observed in relatively turbid environments. Spectra for classes 2 and 3 show high values from 490 to 555 nm while r_n for classes 4 to 7 are mostly characterized by a peak at 490 nm. Conversely, classes 8 to 16 exhibit a decreasing spectrum from the blue to the red wavelengths more typical of relatively clear waters. For all classes, the variations within a class are small in terms of normalized spectra r_n, while the variations are much higher for the raw spectra as the amplitude of R_{RS} is influenced by the gradient of concentrations of optically significant constituents (Fig. 4). In the range of middle wavelengths (490 to 555 nm), class 1 is associated with the highest average amplitude of R_{RS}, followed by class 4. The variations within a class in terms of R_{RS} amplitude decrease with class number (i.e., for more oligotrophic waters) and are very small for classes 15 and 16. The relation between amplitude and shape is discussed further in Section 4.2.

A comparison exercise was performed between the average r_n spectral shapes obtained from the different clusters, i.e., the class centers, and those derived from the Case-1 reflectance model proposed by Morel and Maritorena (2001) in which all optical properties depend on Chl*a* (Morel & Prieur, 1977). For each class the spectrum closest to the class center (in terms of root-mean-square difference) was determined by varying the Chl*a* value. Spectra for classes 9 to 16 can be well represented by the Case-1 model (Fig. 3b) considering oligotrophic to mesotrophic conditions (Chl$a < 0.4\ \mu g\ l^{-1}$) suggesting that Chl*a* (and associated degradation products) can be considered as a major descriptor of the variations in the optical conditions of the corresponding water masses. Conversely, the spectral shapes from classes 8 to 1 increasingly depart from the reflectance spectral shapes allowed by the Case-1 model, particularly in the blue (Fig. 3b). This emphasizes the greater optical complexity of these waters influenced by varying and independent proportions of particulate and dissolved material.

3.2. Average spatial distribution of the classes

An important result is that the set of selected optical water types allows a satisfactory classification of all water bodies in the domain, i.e., with most regions having a total class membership around 1 (see map in supplementary material, Fig. S1). The domain average of the total class membership is 1.14 (s.d., 0.29). It is again highlighted that the classification scheme does not prevent sums of memberships

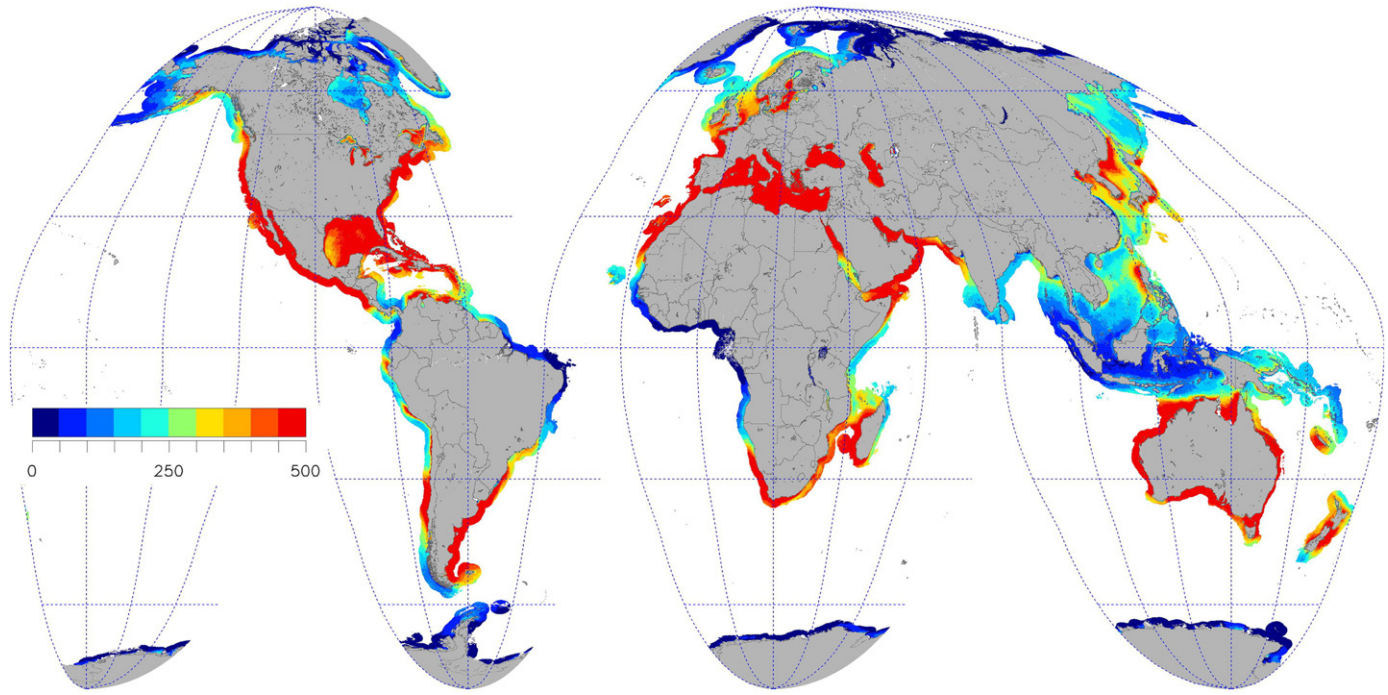

Fig. 2. Number of days with valid data for the period of analysis 1998–2004.

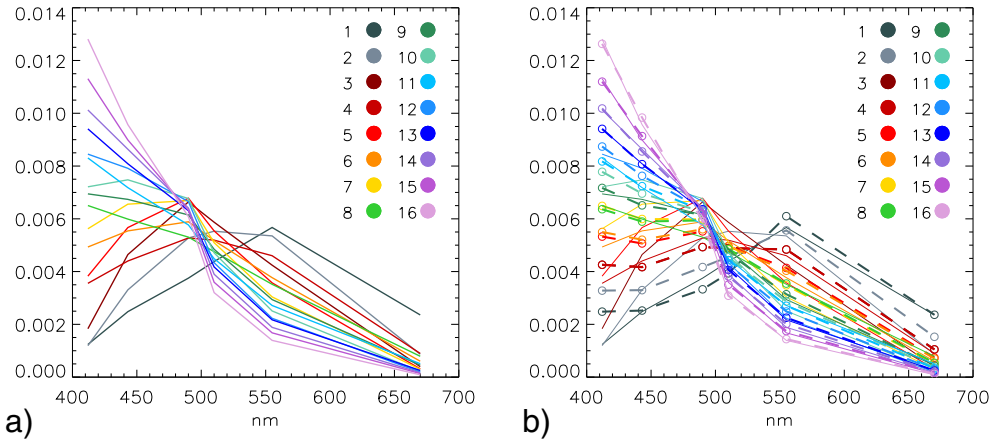

Fig. 3. a) Average (normalized) spectra for the 16 classes. Colors will be used for all figures illustrating class-dependent results. b) The same spectra with spectra obtained with a Case-1 water model (see text) represented by circles and dashed lines.

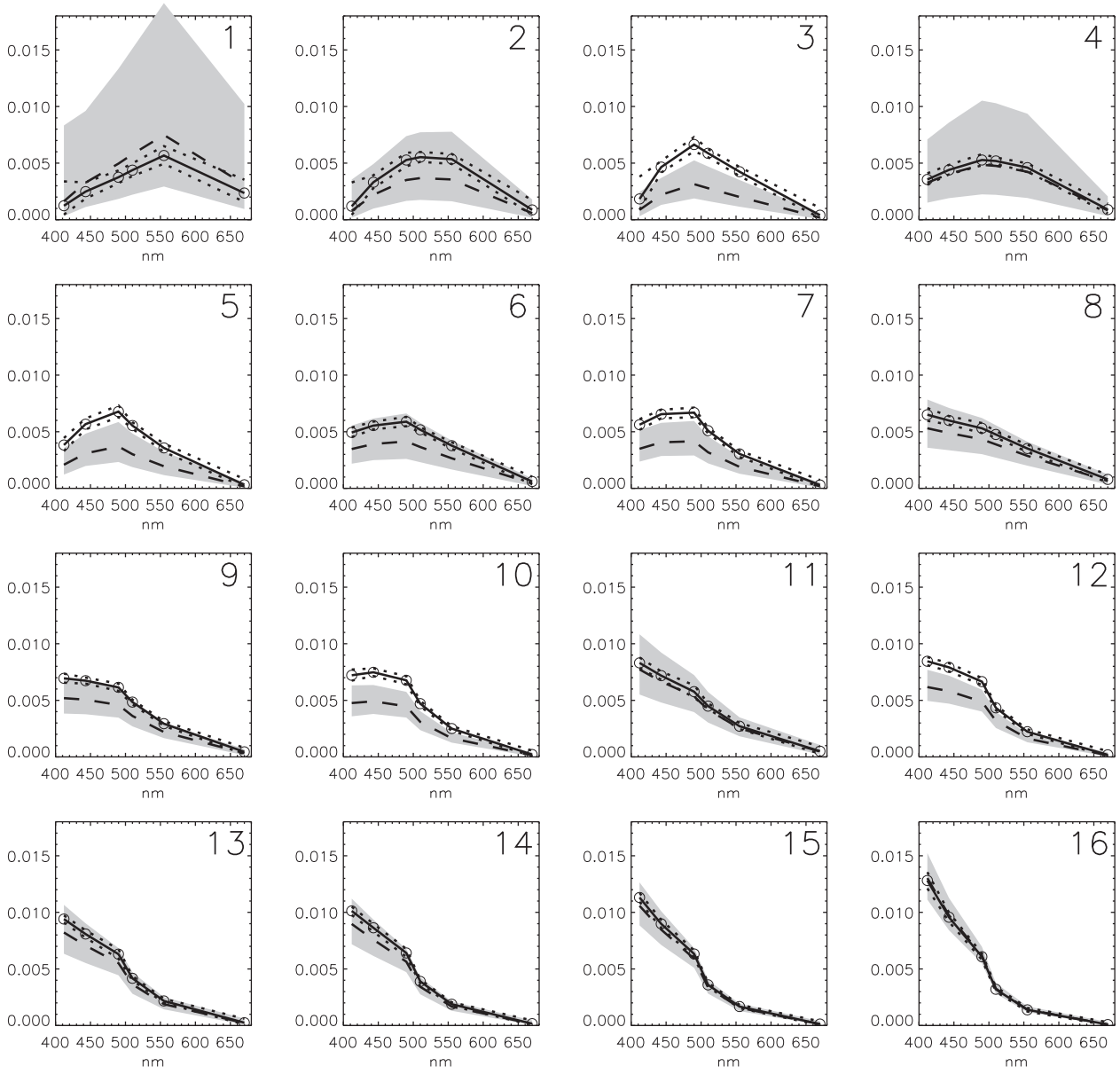

Fig. 4. Average (±1 standard deviation, s.d.) of the reflectance spectra for the 16 classes. Dashed lines (gray envelops) are associated with the average (±1 s.d.) raw reflectance spectra R_{RS}, while lines with circles (dotted lines) are associated with the average (±1 s.d.) normalized spectra. Units are sr^{-1} for the raw reflectance spectra $R_{RS,}$ and nm^{-1} for the normalized spectra.

different from 1: input R_{RS} spectra can be found among class centers and close enough to them to have a cumulative class membership exceeding 1, while a spectrum very different from any class centers may obtain a low cumulative class membership. Some regions usually associated with low data coverage, show total class memberships significantly lower than 1, approximately 0.55 for the Kara and Laptev Seas, or 0.74 for the Antarctic region. The Mediterranean Sea has also a relatively low average total class membership (0.75) due to values around 0.5 in the eastern basin. Waters in this region are ultra-oligotrophic and in general this basin shows some peculiar optical properties with respect to other clear waters (e.g., Claustre et al., 2002; Morel & Gentili, 2009), which might explain why their characteristics have not been fully captured by the classification scheme. On average globally, 64% of the days with valid data are associated with a total membership exceeding 0.9, while this percentage is 80% for a total membership of 0.5.

How is the classification affected by the number of classes? The ISODATA technique was run with parameters of the clustering process modified so as to obtain 12 and 20 classes (i.e., ± 4 with respect to the reference case), in order to test the impact of a coarser and finer classification, respectively. The same classification processing was then repeated on the SeaWiFS data with the case of 12 and 20 classes: the total class membership has global averages equal to 1.01 (s.d., 0.25) and 1.05 (s.d., 0.28), respectively, and shows similar spatial variations as the reference case (not shown). The total class membership thus appears relatively stable with respect to the number of classes, at least in the considered interval. The issue of the number of classes is further discussed in Section 4.2.

Maps of average class membership were created for each class (see supplementary material), while the class most frequently selected as the class of maximum membership (dominant class) is shown in Fig. 5. This indicator is also provided for each region in Tables 1–7 as well as illustrated for selected regions in Fig. 6. As anticipated, the tabulated values are to be treated with caution since they do not reflect internal spatial variations. For instance, the Baltic Sea appears dominated by classes 1 and 2 for its northern and southern parts, respectively, and many shelf regions show a gradient of different dominant classes when going away from shore (examples in Fig. 6). In Fig. 5, the most widespread dominant class is class 14, with 10.9% of the domain, whereas the class least dominant is class 9 (2.2%). Classes 9–16, whose spectral shape can be reasonably well reproduced by a Case-1 water model (Section 3.1), account for approximately half of the domain (48.5%). The corollary is that half of the domain is not well reproduced by such a model and needs another framework to be treated appropriately. It is anyway stressed that providing these statistics does not entail a clear-cut separation between Case-1 and Case-2 waters, which in any case cannot describe all the optical variations expressed by the 16 classes. The statistics also depend on the boundaries of the domain.

High membership for class 1 is found in very coastal waters (Fig. S2) particularly when influenced by a large river outflow. Indicating in brackets the related river, high membership is found in the Eastern Bering Sea (Yukon) and Beaufort Sea (Mackenzie), the Kara Sea (Ob, Yenisei) and the Eastern Siberian Sea (Kolyma). In North America, class 1 is noted in James Bay in the southern Hudson Bay, and in the Lake Erie (but not in the other Great Lakes) and in isolated embayments such as the Chesapeake and Delaware Bays and Long Island Sound (Fig. 6c). In the Gulf of Mexico, class 1 is dominant along the northwestern shore, particularly around the Atchafalaya basin and the Mississippi delta, eastward to Mobile Bay (Fig. 6b). In South America, the outflow of major rivers like the Orinoco, the Amazon and the Rio de la Plata are well detected. In European seas, high values are found prominent in the Baltic Sea, and detected in the North Sea coastal regions including the Thames plume (Figs. 6a and 7), the northwest Black Sea (Danube), the Azov Sea (Don) and the northern Caspian Sea (Volga). Smaller rivers are also picked up by class 1, such as the Gironde estuary, the Loire, Seine and Somme rivers along the French Atlantic coast (Fig. 6a). Northern Indian Ocean regions affected by large rivers are also seen (Indus, Narmada, Ganges, Irrawaddy; see Fig. 6d for the latter two). High membership also affects the Bohai Sea (Huang He or Yellow River, and Liao He) and China Sea (Chang Jiang) and is also found in the Mekong Delta. Class 1 is also found dominant in thin strips along the coasts of some upwelling regions (Benguela and Peru, Figs. 5 and S2).

The distribution of high membership for class 2 is also found in very coastal regions, often in common with class 1 but with the clear difference that the river outflow regions previously cited show a low membership for class 2 (Fig. S3). On the other hand, the Saint-Lawrence estuary is associated with high membership values for class 2, as is the

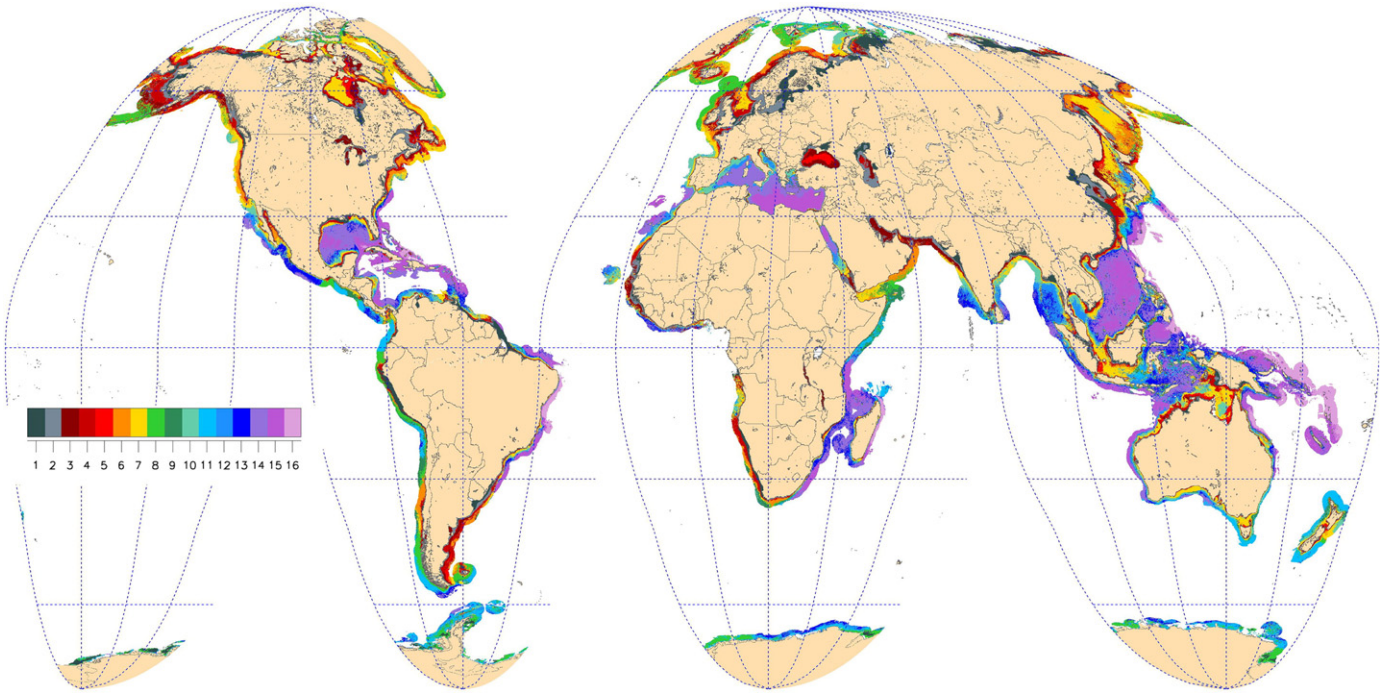

Fig. 5. Class most frequently selected as the class of maximum membership (dominant class) over the period 1998–2004.

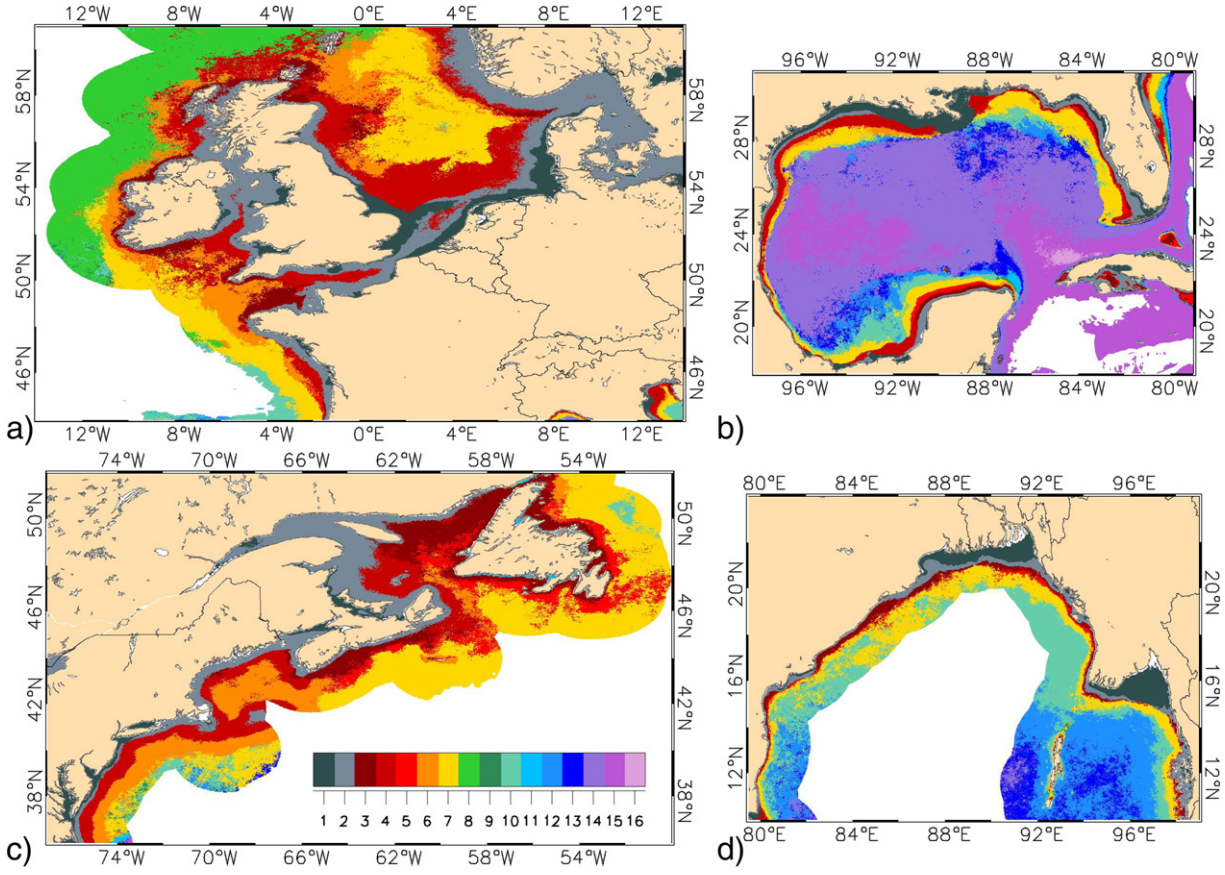

Fig. 6. Class most frequently selected as the class of maximum membership (dominant class) over the period 1998–2004 for selected regions, a) the Northwest European shelf (latitude 44°N–61°N, longitude 14°W–14°E), b) the Gulf of Mexico (18°N–31°N, 98°W–79°W), c) the Northeast American shelf (36°N–52°N, 77°W–50°W), and d) the Bay of Bengal (10°N–24°N, 79.5°E–99°E).

coastal Gulf of Maine (Fig. 6b). Class 2 is fairly dominant along the European Northwest shelf coasts and in the Irish Sea (Figs. 6a and 7). Baltic and Caspian seas are also affected by class 2 but more in their central and southern parts. High membership is found for class 3 in some regions of the Canadian archipelago, in the lakes Huron, Michigan and Superior, and in the Persian Gulf. For the American Great Lakes and African Lakes as a whole, and for the Black and Caspian seas, this is the class with the most frequent maximum class membership (Tables 2 and 7). Class 4 shows a strong signal slightly offshore in the Eastern Bering Sea and the northern part of the East China Sea. As for previous classes, high membership values are detected in some parts of the Indonesian archipelago and the northern Australia coastal waters.

High-numbered classes do not show features as distinct as the previous ones, even though some exceptions can be noticed. For instance, class 5 appears strongly in the Black Sea and eastern Hudson Bay, while the western Hudson Bay is dominated by class 7, which is the class of most frequent maximum membership for the Bay as a whole (Figs. S6 and S8, Table 1). The Andaman Sea is most clearly associated with class 12 (Figs. 6d and S13). Otherwise, patterns are less marked and take often the form of gradients. Class 4 shows patterns of high membership in the Irish Sea, the English Channel and the southern North Sea (but detached from the coasts), while classes 5 and 6 appear more on the Atlantic shelf and the northern North Sea. Finally, class 7 is found at the northernmost part of the North Sea and close to the Atlantic shelf slope while class 8 appears in the outer Atlantic waters (Figs. 7, and S5 to S9).

The most oligotrophic waters are represented by classes 14, 15 and 16 (Figs. S15 to S17). The Western and Eastern Mediterranean waters are mostly related to classes 14 and 15, respectively. These two classes are also important for the central Gulf of Mexico (Fig. 6b) and the Caribbean archipelago, the eastern Brazilian coast, the Mozambique Channel and the South China Sea. Finally, class 16 is observed along the eastern Brazilian coast and the western Pacific edge.

From the multi-annual average distribution, the class memberships have been associated with various intervals of water depth and distance from the nearest coast, and the resulting histograms are shown in Fig. 8. Classes 1 to 4 tend to have high membership values close to the coast and for shallow waters. For a distance to the coast beyond 50–75 km, the histograms do not show large variations, but beyond 200 km, the membership to classes 13 to 16 increases markedly. Between 75 and 400-m water depth, the membership for low-numbered classes tends to decrease while the opposite is seen for high-numbered classes. Beyond 450-m depth, there is an increase in membership for classes 14 to 16.

3.3. Similarity between regions

One of the goals of the study is to identify regions that are optically similar. This can be done by examining the maps of average class memberships for all classes (Figs. S2 to S17) which is cumbersome to reach synthetic conclusions. To support that process, the average class membership (16 values) for all regions was input to a hierarchical clustering (Euclidian distance, Ward's algorithm) to illustrate the respective similarities between regions (see dendrogram in Fig. S18). The aim of this procedure was merely to help in finding general similarities between regions and to build Fig. 9. It is also well understood that an analysis based on regions depends on their boundaries and hides some of the variations existing within each region.

Regions most influenced by oligotrophic waters are grouped together (Fig. 9a). The average class membership, noted $<p>$, has its maximum

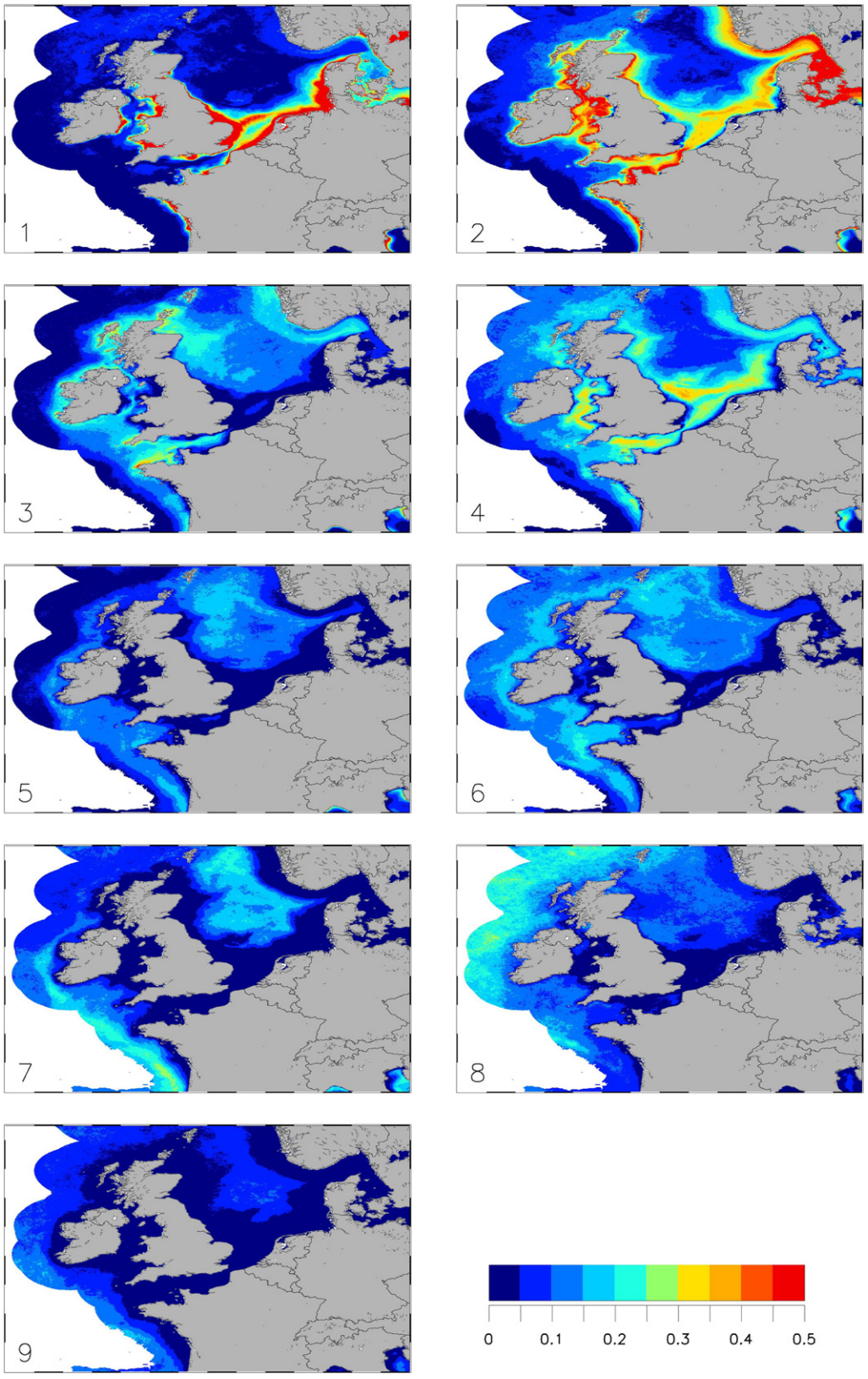

Fig. 7. Average (normalized) class membership for classes 1 to 9 associated with the Northwest European shelf (latitude 44°N–61°N, longitude 14°W–14°E).

between classes 14 and 16 (the East Brazil and Western Pacific regions). Before class 9, most <*p*> values are very low, except for a secondary peak for class 7. The Mediterranean Sea shows marked secondary peaks for classes 10 and 12 that can be related with various patterns within the basin. Oligotrophic waters are dominant for other regions as well but with a stronger contribution from other water types (Fig. 9b). For instance, <*p*> is approximately 0.04–0.05 for the Gulf of Mexico for the low-numbered classes, indicating the significant impact

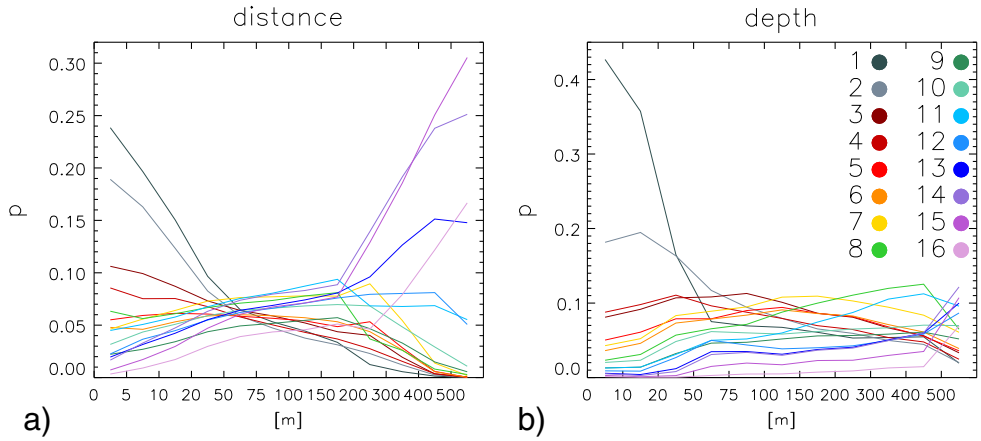

Fig. 8. Average class membership as a function of a) distance from the coast, and b) water depth.

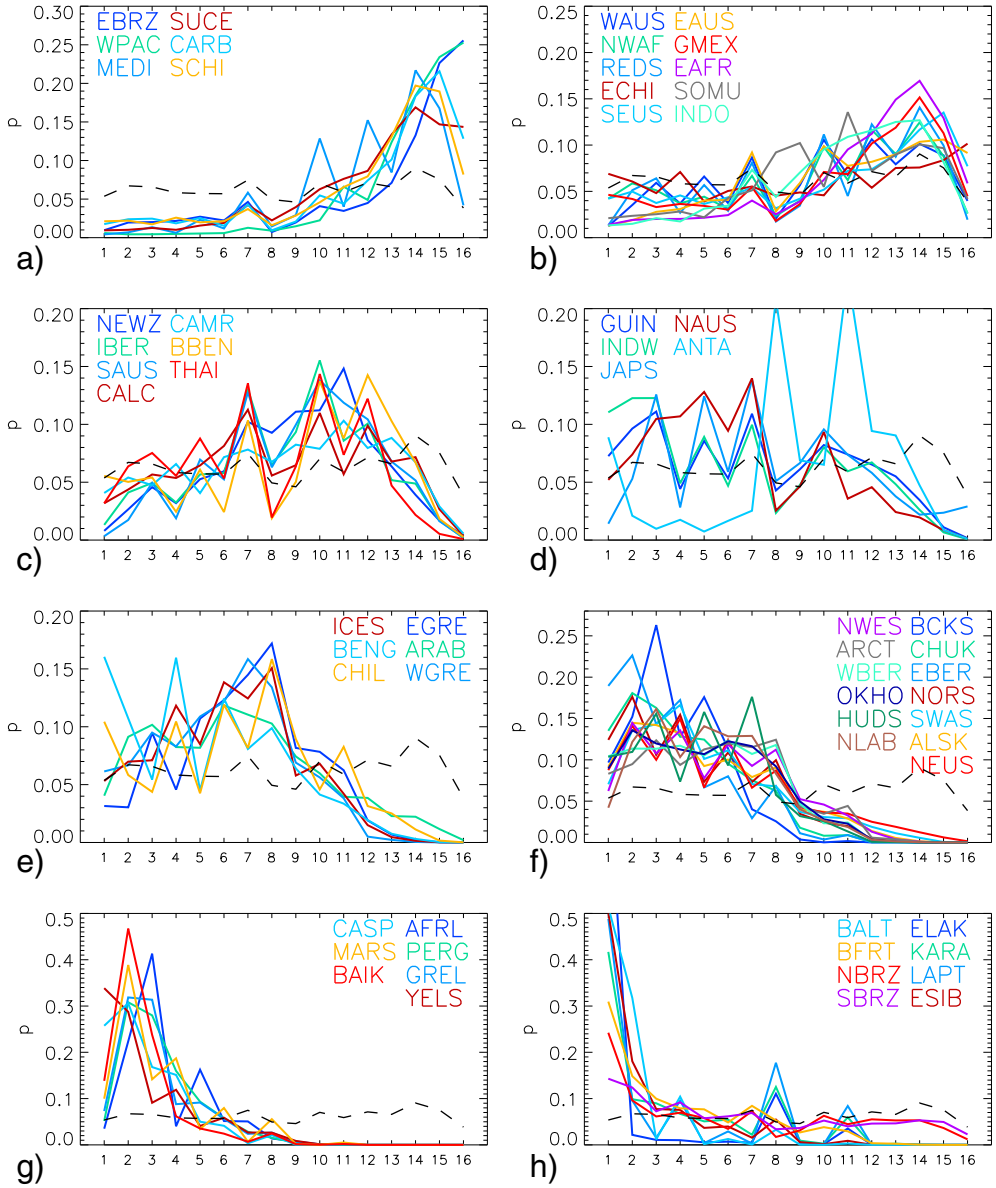

Fig. 9. Average class memberships for all selected regions. The dashed black line represents the global average on each graph.

of the coastal waters and river outflows in that basin. All the regions in these two groups (Fig. 9a,b) tend to be associated with fairly low latitudes, the Mediterranean Sea being the most poleward region.

Some regions, most often located in low to mid-latitudes, have an average class membership distributed across many or all classes (Fig. 9c,d). Some of the major eastern boundary upwelling regions are found in these groups, the California Current region (CALC) and the Iberian upwelling (IBER). In Fig. 9c, regions have a maximum <*p*> for classes between 10 and 12 (except CALC with three similar peaks for classes 7, 9 and 12) with many secondary peaks. For other regions, the <*p*> distributions tend to be more shifted towards low-numbered classes (Fig. 9d) and show various peaks for classes 3, 5, 7 and 10. Following its position in the dendrogram (Fig. S18), the Antarctic coastal region (ANTA) is included in Fig. 6d, and show two peaks for classes 8 and 11 and a secondary peak for class 1.

Fig. 9e shows regions, found at various latitudes, with low membership beyond class 9 and with peak <*p*> values observed between classes 6 and 8, except the Benguela current region (BENG) that shows a high <*p*> value for class 4 and class 1, highlighting the importance of turbid waters along a thin coastal strip. The Peru/Chile (CHIL) upwelling also shows a (secondary) peak for class 1. Other regions display a further shift towards low-numbered classes, with <*p*> maxima mostly between classes 2 and 4 (Fig. 9f) and low memberships beyond class 8. The North Atlantic continental shelf regions (Northeast US shelf, NEUS; Northwest European shelf, NWES; and North Sea, NORS) all show peaks for classes 2, 4, 6 and 8. The Black Sea stands out with a clear <*p*> maximum at class 3; included in that region, the Azov Sea is dominated by class 1 (Fig. S2). High latitude regions, like the Arctic area (ARCT), the Western Bering Sea (WBER), the Hudson Bay (HUDS) and the Newfoundland-Labrador (NLAB) show primary or secondary peaks for classes 5 to 8.

Some marginal seas and inland water bodies are grouped together in Fig. 9g, displaying low memberships beyond class 5 and maxima found for classes 2 and 3. Even in these fairly small water bodies, there is a significant spatial variability. For instance, the Lake Erie part of the American Great Lakes (GREL) shows a dominance of class 1 (Fig. S2) which is not visible for the region average. The <*p*> values for class 1 are fairly low, except for the Caspian Sea (CASP) and Yellow Sea (YELS) that are found close to each other in the dendrogram (Fig. S18). On the contrary, class 1 displays the maximum <*p*> values for all regions grouped in Fig. 9h while high-numbered classes are associated with low memberships, except for some secondary peaks for classes 8 or 11. Besides the Baltic Sea (BALT) and the European Lakes (ELAK), they are all associated with a large river system (see Section 3.2).

The analysis of Fig. 9 hints at a latitudinal dependence of the average class membership. This is confirmed by Fig. 10a that shows a clear dominance of classes 9 to 16 (amounting to a cumulated membership of approximately 0.7) for latitudes between 40°N and 30°S. On the contrary, the northern mid to high latitudes (40°N to 75°N) are mostly associated with classes 1 to 8 (amounting to a cumulated membership of approximately 0.8). This is also true to a lesser extent in the southern hemisphere (from 30°S to 55°S) that has much less land mass in its latitudinal band.

3.4. Optical diversity

In this section, the local variations of optical conditions or status are investigated. For any given day and bin, few classes are needed to explain most of the total class membership: on average globally, 2.2 and 2.5 classes account for 90% and 95% of the (normalized) total membership, respectively. But for any given bin there is a significant variability in optical properties in time: for instance, on average globally, 10.0 classes are selected at least once as the class of maximum membership. Focusing just on the main patterns, considering the dominant classes accounting for 90% of the days with valid data, 5.2 different classes are on average selected as dominant. A budget focusing on the occurrence of the dominant class is however ignoring the role of secondary classes, albeit they might concur to a large part of the total membership. A more comprehensive indicator of optical diversity is introduced by adopting the Shannon index (Shannon, 1948) often used in biodiversity studies:

$$H = -\sum_{ic=1}^{NC} p(ic)\,\ln[p(ic)] \tag{6}$$

where *NC* is the number of classes represented for a given bin. As we are interested in the overall diversity manifested in time at a given place, this calculation is performed from the multi-annual average map of class memberships. Then regional averages are computed and listed in Tables 1–7. To give some background to this analysis, it is recalled that *H* reaches a maximum value equal to the natural logarithm of *NC*, $\ln(16) = 2.8$, if the 16 classes have occurred with the same frequency, which is the case of maximum diversity. A case with only 2 classes equally represented leads to $H = 0.7$.

Fig. 11 shows that *H* is mostly between 1.5 and 2.8, with a global average of 1.98 (s.d., 0.48). There are however regions where *H* is below 1.5. This is particularly the case in the Baltic Sea and the southern North Sea, and in coastal regions close to river outflows, where class 1 is often dominant (compares Figs. 5 and 11). For instance the Yellow Sea (YELS) has an overall index of 1.37 (Table 5). Some of the regions with low *H* are found in the Arctic: KARA, LAPT, ESIB have an average *H* around 1, and CHUK and BFRT's *H* is approximately 1.6–1.9 (Table 1). The coastal Antarctic waters show a diversity of 1.82.

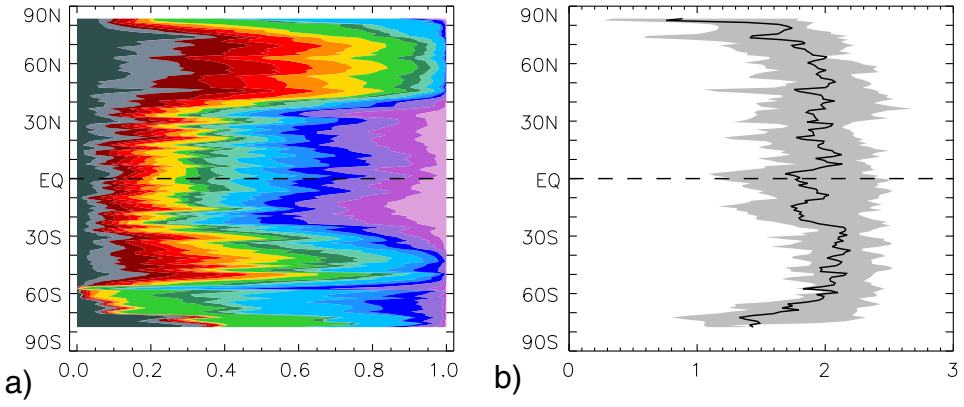

Fig. 10. Latitudinal dependence of a) the average class membership, and b) the Shannon index. In a), the colors are those associated with the classes in Fig. 3. In b), the line indicates the average while the gray area shows ±1 standard deviation.

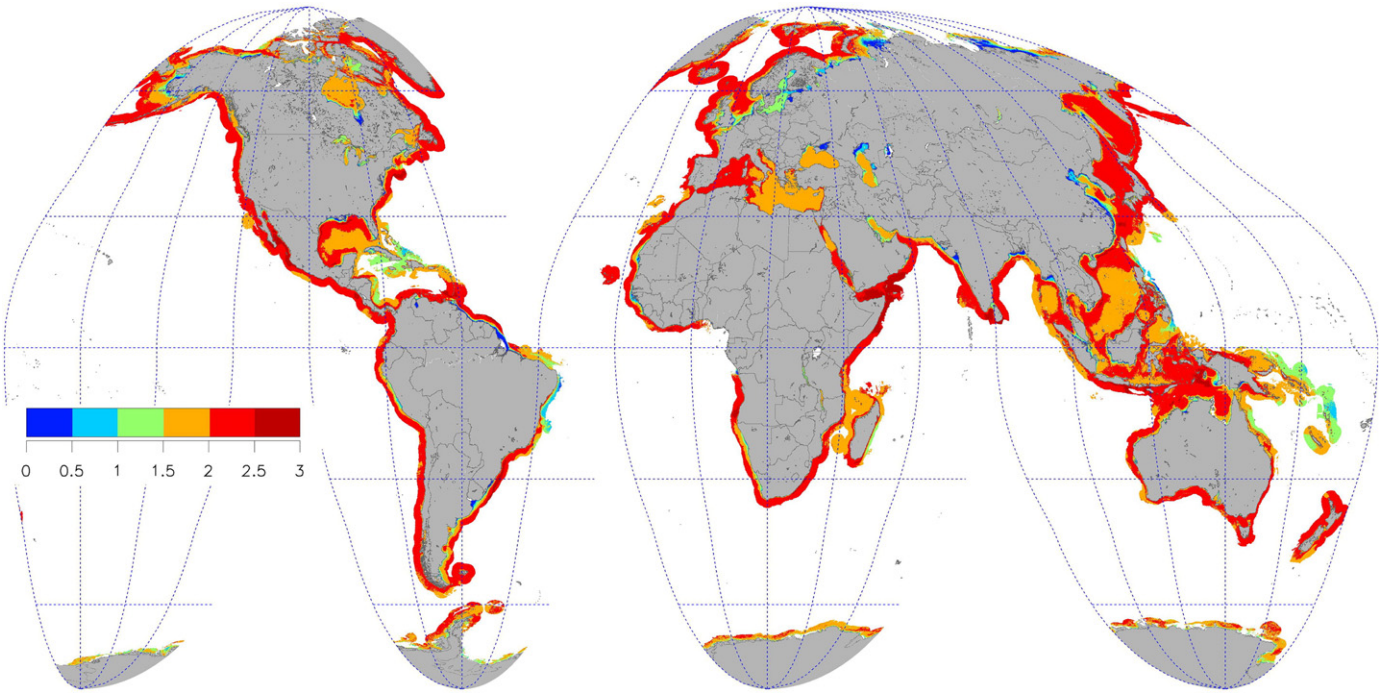

Fig. 11. Shannon index computed from the multi-annual class memberships.

Other regions with a relatively low diversity are associated with clear water classes as found in the Caribbean, the East Brazil region (EBRZ, with an average H of 1.53) and the West Pacific fringe. Closed seas also tend to have a relatively low average index, approximately 1.6–1.7 for the Black and Marmara Seas or the Persian Gulf, while inland water bodies have an even lower index (between 0.6 and 1.5, Table 7).

Regions of high diversity are often found in intermediate regions between coastal domain and open ocean. For instance, H is observed to increase from less than 0.5 close to the Mississippi river mouth to more than 2.5 in an offshore strip, to decrease again to less than 2 in the middle of the Gulf of Mexico (Fig. 11). This is translated in Fig. 8a by a relative convergence of class memberships in the interval of distance-to-coast of 50–75 km.

The number of classes has a direct impact on the optical diversity: the more classes are considered, the higher H is likely to be. The results shown for H should be interpreted in terms of relative distribution of diversity, i.e., describing regions that appear more or less optically diverse compared to others. The impact of the number of classes can be verified with the classification based on 12 and 20 classes. The same calculations described above lead to a multi-annual global average of 1.75 (s.d., 0.50) and 2.21 (s.d., 0.60) for the cases of 12 and 20 classes, respectively. But the spatial distribution of H remains virtually unchanged (not shown) which lends a general validity to the results on optical diversity discussed in this paper.

3.5. Illustrations of the temporal dynamics

All statistics presented so far had a focus on spatial variability. Obviously, the multi-annual averages are modulated by annual cycles specific to each region that affect the local optically significant constituents. These variations in turn are reflected in the distribution of class memberships. A comprehensive description of seasonal variations associated with the optical classification is not undertaken here for the sake of space, but representative examples are provided to illustrate how local annual cycles are captured in terms of relative class memberships.

First, monthly averages of the class memberships are computed, from which a monthly climatology is derived. The average climatology is presented for four regions in Fig. 12.

The climatology obtained for the Kara Sea (KARA, Fig. 1) is restricted to the period May to September (Fig. 12a), as ocean color observations are limited in this region by ice and cloud cover and low solar zenith angles. Over the year, the dominant class is class 1 (Table 1) but classes 8 and 11 have larger memberships in spring, while classes 1 to 6 account for a cumulated membership of 0.8 in August–September. Siberian rivers are large providers of dissolved organic matter (Opsahl, Benner, & Amon, 1999) and secondarily of sediments (Gordeev, 2006). These fluxes increase from May to July with the thaw (Le Fouest, Babin, & Tremblay, 2013). Erosion from permafrost coastlines might also contribute an influx of material on the shelf (Lantuit et al., 2012). Added to this accumulation of dissolved matter and particles, the spring/summer biological activity contributes to a shift towards turbid conditions reflected in the shift in dominant classes. For these Arctic regions, a word of caution is however warranted. The total class membership found in the Kara sea is fairly low (Fig. S1), which may be partly explained by a small contribution to the training data set. The R_{RS} values observed in these absorbing waters (not shown) are often low and affected by a significant level of noise. Retrieval conditions for the atmospheric correction are as well challenging, with low zenith angles and the impact of sea ice (e.g., Bélanger, Ehn and Babin, 2007).

A second example is the climatology observed in the northwest Mediterranean Sea (latitude 41°N–44.5°N, longitude 5.45–9.4°E). In summer, the area is characterized by mostly oligotrophic waters, with classes 9 to 16 accounting for almost the entire class membership (Fig. 12b). The phytoplankton annual cycle in the region shows a strong bloom in spring (e.g., D'Ortenzio & Ribera d'Alcalà, 2009; Marty, Chiavérini, Pizay, & Avril, 2002), which is reflected by an increased class membership by the low-numbered classes; classes 2 to 7 thus account for more than half of the total membership in March.

In Fig. 12c is shown an example of coastal regions associated with upwelling conditions, the Iberian region (IBER). The contribution of the low-numbered classes peaks in March, with a cumulative class membership for classes 1 to 7 reaching 0.6. Upwelling favorable winds can occur year-long but are more frequent in summer. Averaged over the entire region, the annual cycle of class membership appears affected by phytoplankton spring bloom conditions found along the western coast (Joint et al., 2002) or in peripheral areas such as the Gulf of

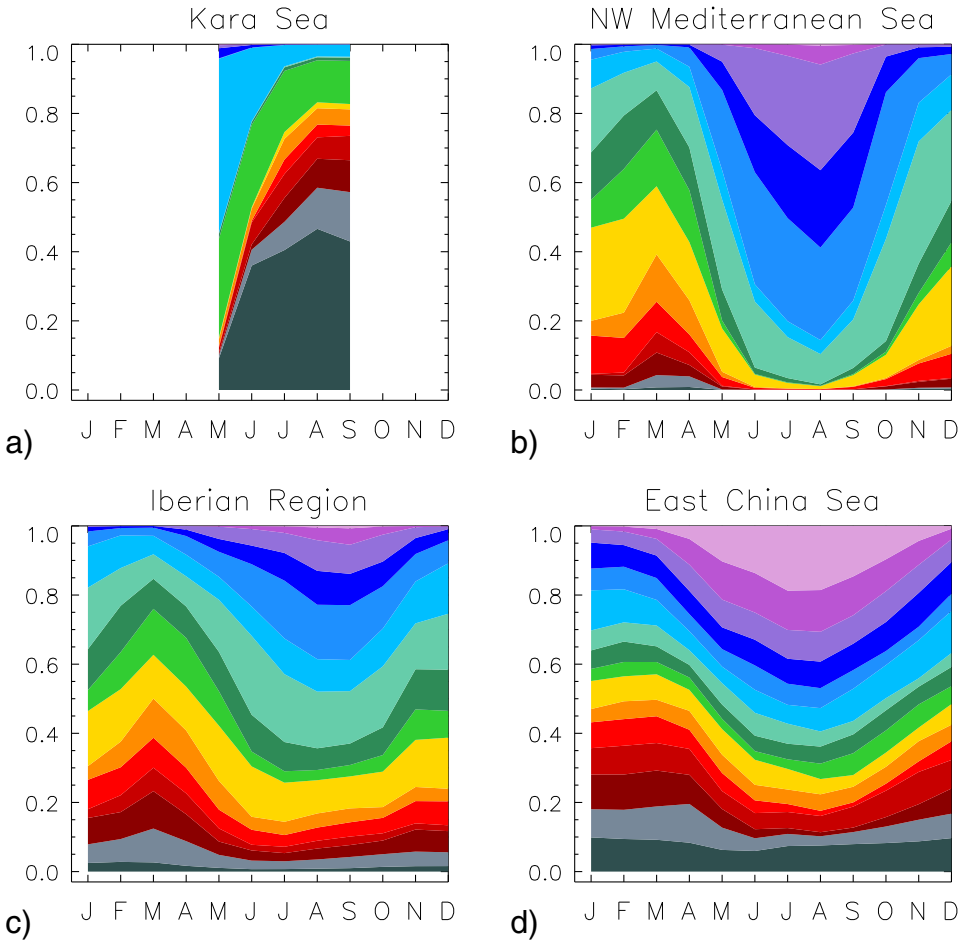

Fig. 12. Monthly climatology of the class membership for selected regions, a) the Kara Sea, b) the northwest Mediterranean Sea (latitude 41°N–44.5°N, longitude 5.45°E–9.4°E), c) the Iberian coastal region and d) the East China Sea. Colors are as in Fig. 3.

Cádiz (Navarro & Ruiz, 2006) or the Cantabrian shelf (Álvarez, Nogueira, Acuña, López-Álvarez and Sostres, 2009).

The East China Sea (ECHI, Fig. 1) also displays a clear annual cycle (Fig. 12d). Classes 1 and 2, dominant in coastal waters (Fig. 5) maintain a combined membership between 0.1 and 0.2 through the year. Classes 1 to 7 account for a cumulated membership of approximately 0.5 in winter and early spring, while clearer waters gain more importance at the end of the summer. A similar phase is observed for the Yellow Sea/Bohai Sea region (YELS, Fig. 1) even though in that case the low-numbered classes account for most of the membership (not shown). Discharge of the major rivers in these areas actually peak in summer (Ning, Liu, Cai, Fang, & Chai, 1998; Wang, Yang, Saito, Liu, & Sun, 2006) but their direct influence may be limited by damming and may not extend far offshore, while sediment resuspension by wind is a strong contributor to the regional turbidity in winter (Shi & Wang, 2012). Phytoplankton blooms in spring/summer (Gong, Wen, Wang, & Liu, 2003; Zhou, Xuan, Huang, Liu, & Sun, 2013) are another important element of a complex set of factors affecting optical variability in the region.

4. Discussion

4.1. Optical classification and remote sensing

A first element of discussion is about the potential and limitations of remote sensing for optical classification. For global applications, the potential allowed by remote sensing since 1997 in terms of spatial, temporal and spectral resolution is that of SeaWiFS-like missions, and the optical variability that can be detected is tied to the spatial resolution of the sensor. For this study, a grid with a resolution of 1/48th-degree (approximately 2.3 km) was selected using LAC SeaWiFS data. This is to our knowledge the first time LAC SeaWiFS data are used systematically at global scale, albeit for coastal applications (see Mélin, Steinich, Gobron, Pinty and Verstraete, 2002, for terrestrial vegetation science) and this resolution is at least twice that of standard global products. SeaWiFS pixel size is approximately 1.1 km at nadir and increasing away from nadir, so that the selected grid together with the daily sampling is as close as possible to pixel-size information. In that context, 1-km pixels should be able to capture most reflectance spectral characteristics. Assuming that temporal variations at a fixed station may represent small-scale spatial variations around that location, some studies based on time series suggest that local heterogeneity is not translated in large variations in optical properties. For instance, variations in R_{RS} observed at a coastal station in a time interval of 1 to 3 h was approximately 4% (10% in the red; Mélin & Zibordi, 2007; Zibordi, Mélin, & Berthon, 2006). On the other hand, significant variations along spatial scales of tens of meters have been documented for extreme events such as cyanobacteria blooms (Kutser, 2004). It is therefore acknowledged that fine-scale features with specific spectral signatures might be blurred within a 1-km spatial scale. If these specific features are associated only with small patches (significantly smaller than 1 km) or occur only rarely in larger dimensions, they are unlikely to be present in the training data set and are not well covered by the proposed classification. Finally, it cannot be excluded that by chance specific water types have not been included in the training data set, or only in minute amounts, notwithstanding its size.

Another limitation of the proposed classification is associated with its spectral resolution restricted to six bands between 412 and 670 nm. In highly turbid waters found in very coastal or estuarine waters, optical variability can be pronounced for wavelengths longer than 670 nm (e.g., Doxaran, Cherukuru, & Lavender, 2006; Lubac & Loisel, 2007), and cannot be captured by a sensor like SeaWiFS. The addition of a band at 709 nm such as found with the European Space Agency Medium Resolution Imaging Spectrometer (MERIS) or on its upcoming Ocean and Land Colour Instrument (OLCI) would be precious in that regard. Eventually, the results presented in this study are to be interpreted in the context of standard mid-resolution satellite ocean color imagery.

4.2. Clustering and classification approach

The classification developed here relies on normalized reflectance spectra, i.e., on the R_{RS} spectral shape, which deserves some comments. First, in natural waters, R_{RS} shape and amplitude are not independent (which partly explains the success of band-ratio algorithms) while the different R_{RS} bands are often well correlated (Lee, Shang, Hu, & Zibordi, 2014). In the context of a Case-1 water model (Morel & Maritorena, 2001), shape and amplitude are even equivalent so that classifications with or without normalization would give similar distributions (note that the dispersion around the mean decreases for the high-numbered, meso- to oligotrophic, waters with respect to turbid waters, Fig. 4). This being said, both shape and amplitude bring fruitful information. While recognizing that all optically significant constituents have an effect on both amplitude and shape of R_{RS} and that the relationship between the two components may not be unique (e.g., Sydor, Gould, Arnone, Haltrin, & Goode, 2004), a large part of the variance in R_{RS} (without normalization) is associated with back-scattering or to the concentration of (mainly non-chlorophyllous) particles, whereas absorption terms (by phytoplankton cells, CDOM or detritus) have a relatively larger impact on the spectral shape (Loisel & Morel, 2001; Sathyendranath, Prieur, & Morel, 1989). Consequently, by normalizing the R_{RS} spectra, the choice is made to give preference to the latter in the classification process, whereas in the absence of normalization, the resulting distribution of classes is much influenced by a gradient in concentrations, particularly of particles, in coastal waters. Taking the example of a water type strongly dominated by non-algal particles like sediments (typically represented by class 1), and assuming that this dominance is associated with a certain spectral shape, spectra associated with significantly different concentrations of sediments end up as members of the same class if the normalization is employed, which is a welcome feature for the application we envisage. Moreover, the distribution of classes obtained here met our expectations in terms of spatial and temporal distributions (Figs. 5, 12, and S2 to S17).

The choice of approach is ultimately dependent on the application. To some extent, a more dynamic range of concentrations is allowed within a class when normalization is employed, which might be a positive feature in the context of bio-optical algorithms development and application. On the other hand, a class-based approach aiming at detecting thresholds would not use a prior normalization. Full spectra have also been used to display multispectral field measurements onto a two-dimensional representation (e.g., Zibordi, Berthon, Mélin, & D'Alimonte, 2011) In any case, more studies are due to fully comprehend the advantages specific to each approach.

For a given set of parameters, the ISODATA unsupervised clustering technique yielded 16 classes. The details of the classification results change with the number of classes but the main patterns are conserved, at least if the number of classes allows a sufficient distinction between different water types (tests were performed with 12 and 20 classes). In that respect, the number of classes cannot be too low, while it should not be too high as to become difficult to manage. Metrics exist to determine the optimal number of classes from a mathematical point of view (Moore et al., 2009), and are readily applied to cases presenting clustered distributions associated with specific environments often seen with collection of field measurements (e.g., Zibordi et al., 2011). This is less true when considering a continuum of conditions (in the N-dimension space of reflectance, where N is the number of bands) such as is found with a global satellite data set. The present selection of 16 classes (half of which can be interpreted in a Case-1 water framework — Section 3.1) does not pretend to represent a mathematical optimum but is rather a heuristic choice that serves our purpose well: producing a discrimination of the water masses useful to discuss their respective distributions while achieving a comprehensive classification of the waters studied. The last point was particularly successful (as quantified by the total class membership), even though it is recognized that some areas are still imperfectly classified. This is the case for waters with very low reflectance values that, being very noisy, likely will be hard to classify with the current approach regardless of the number of classes. In operational applications, a specific status for these cases could be envisioned. The Eastern Mediterranean basin also shows relatively low total class memberships while that area is fairly well represented in the training data set that has been specifically constructed to span as well as possible all the optical variability found in natural waters. It is possible that subtle variations in the associated R_{RS} with respect to other oligotrophic regions (Claustre et al., 2002) be blurred in the overall classification of oligotrophic waters. Therefore, further developments should include particular attention to some regions in the construction of the training data set and the clustering process.

4.3. Applications

The main objective of this study was to document the optical variability observed at global scale for the coastal ocean, to assess similarities between regions and the optical variability within a region. Further applications can be envisioned.

One application is the selection of algorithms for specific regions. Some areas are not well characterized by optical measurements and the capacity to relate these to other regions that are much more documented in terms of optical properties, and for which specific algorithms or bio-optical parameterizations exist, offers the possibility of a more informed choice in that respect. More generally, optical classification can be used to define the range of applicability of a specific algorithm (e.g., D'Alimonte et al., 2003; Mélin et al., 2011) and to better constrain the related uncertainties (Moore et al., 2009). Similarly it is possible to test the representativeness of an in-situ data set collected in a specific region and season with respect to other regions and times of the year. If algorithms are available for a set of classes, their outputs can be combined seamlessly using the class memberships as weighting factors (e.g., D'Alimonte et al., 2003; Mélin et al., 2011; Moore et al., 2009, 2014; Vantrepotte et al., 2012).

Other applications can be envisaged in a context broader than marine optics. The optical classification of coastal waters and marginal seas, and the analysis of similarities between regions, could usefully be combined with factors related to the distribution of benthic communities (e.g., related to light availability, Gattuso et al., 2006) or other classification schemes applied to coastal areas like typologies of a various nature (e.g., Buddemeier, Smith, Swaney, Crossland, & Maxwell, 2008; Dürr et al., 2011). In general, this classification framework could help refine distributions of biogeographic provinces (Longhurst, 2006; Reygondeau et al., 2013).

Ecological applications can be further discussed. The general variations in optical types at any location have been addressed by quantifying the number of classes selected as dominant during the period and an index of optical diversity. In turn, the results on optical diversity can be linked to indices of marine biodiversity. A first general result is that the optical diversity H tends to decrease as latitude increases beyond 60° (Fig. 10b). The dependence with latitude is otherwise fairly small, with the exception of local minima in the equatorial and southern tropical regions, associated with areas like northern South America and the

western Pacific edge (Fig. 11). Even though knowledge on marine biodiversity is still limited, it is interesting to note that the decrease in *H* at high latitudes presents similarities with latitudinal patterns of biodiversity estimates for phytoplankton (Barton, Dutkiewicz, Flierl, Bragg, & Follows, 2010), zooplankton (Rutherford, D'Hondt, & Prell, 1999) and higher trophic levels (Tittensor et al., 2010; Worm, Sandow, Oschlies, Lotze, & Myers, 2005), while this does not necessarily apply to benthos (Brandt et al., 2007). De Monte, Soccodato, Alvain, and d'Ovidio (2013) presented a similar general behavior for satellite-derived estimates of optical diversity for the open ocean.

Optical water types seem to coexist more at intermediate distances from the coast (Fig. 8a) and there are various regions for which *H* increases away from shore before decreasing again (Fig. 11). Considering that on average phytoplankton biomass tends to decrease going away from shore, the distribution of the optical diversity can again be related to results on biodiversity. Indeed, some studies have suggested that aquatic biodiversity peaks at intermediate levels of phytoplankton biomass or productivity (Dodson, Arnott, & Cottingham, 2000; Irigoien, Huisman, & Harris, 2004; Li, 2002; Uusitalo et al., 2013). The multi-annual averages considered in this work also integrate the effect of seasonal and local variations. Regions like upwelling centers tend to show high phytoplankton biomass values very close to the coast and low values far offshore, while the extension of the area with high values fluctuates with the upwelling regime (Carr & Kearns, 2003; Thomas & Strub, 2001), so that optical diversity integrated over the annual cycle displays a peak at intermediate distances from the coast. Optical diversity also tends to increase in some transition zones characterized by frequent occurrences of filaments or frontal structures.

Links between optical diversity and biodiversity estimates are not surprising. Light with its variations in both quantity and spectral character, is a crucial element shaping ecological communities. Irigoien et al. (2004) underlined a relationship between phytoplankton biodiversity and a shading index. Pigment assemblages in cyanobacteria might be an element defining their respective ecological niches (Ting, Rocap, King, & Chisholm, 2002) and competition for light (Huisman, van Oostveen, & Weissing, 1999) is one important pressure selecting dominant algal species. In general, the light climate has a strong impact on phytoplankton communities (e.g., Finkel et al., 2006; Moore et al., 2006; Ragni & Ribera d'Alcalà, 2004; Sathyendranath & Platt, 2007) while in turn underwater irradiance is influenced by algal cells. The light climate exerts pressure on higher trophic levels as well, from zooplankton to fish (Aksnes, Nejstgaard, Soedberg, & Sørnes, 2004; Eiane, Aksnes, Bagøien, & Kaartvedt, 1999; Wissel, Boeing, & Ramcharan, 2003; Zettler & Carter, 1986). Classification schemes can contribute to thoroughly integrate optics into ecological studies.

5. Conclusion

This study has presented a classification scheme focused on coastal regions and marginal seas. By using satellite data and maximizing the extent of the training data set in terms of spatial coverage and annual cycle, the scheme is able to perform a satisfactory classification of all water bodies in the domain. This result holds for regions not included in the training data like inland water bodies. The classification allows the quantification of the optical similarity between regions. Some regions have been better sampled than others by optical measurement programs and, as mentioned as one justification for the work, the ability to find similarities between water bodies is potentially useful to support the application of bio-optical relationships and algorithms related to one region to another region.

The set of 16 classes used in this work covers very turbid waters found close to river outflow regions (class 1) to oligotrophic waters (class 16). Logically, high-numbered classes are observed more often as the distance from the coast increases. This is also true for bathymetry even if the dependence is less clear. Averaged zonally, mid-latitude regions are more affected by turbid waters (low-numbered classes) than subtropical or polar regions. Examples of temporal changes in the class memberships have been shown for specific regions, illustrating specific variations as the annual cycles of optically significant constituents proceed. The general variability in optical types at any location has been addressed by quantifying the number of classes selected as dominant during the period and an index of optical diversity that has been linked to indices of marine biodiversity.

Ultimately, the marine reflectance spectrum is an overall manifestation of ambient conditions in the upper layer of the water column in terms of phytoplankton abundance and type, biotic detritus, minerals and chromophoric dissolved organic matter. An optical classification scheme appears as a powerful diagnostic tool to be applied to aquatic ecosystems in monitoring strategies and ecological studies. The distribution in optical water types brings more comprehensive information than single variables, like chlorophyll-*a* concentration or turbidity, and as discussed above, has interesting connections with fundamental ecological properties like biodiversity.

Acknowledgments

The authors thank the Ocean Biology Processing Group of NASA for the distribution of SeaWiFS data. The British Oceanographic Data Centre is acknowledged for the GEBCO data set. This research has been funded by the GlobCoast project (www.foresea.fr/globcoast) which is funded by the French Agence Nationale de la Recherche (ANR 2011 BS56 018 01). The GlobCoast project is affiliated to the LOICZ and AQUIMER. The study also contributes to the Climate Change Initiative of the European Space Agency. Finally, a reviewer is thanked for useful comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2015.01.023.

References

Aksnes, D.L., Nejstgaard, J., Soedberg, E., & Sørnes, T. (2004). Optical control of fish and zooplankton populations. *Limnology and Oceanography*, *49*, 233–238.
Álvarez, E., Nogueira, E., Acuña, J.L., López-Álvarez, M., & Sostres, J. (2009). Short-term dynamics of late-winter phytoplankton blooms in a temperate ecosystem (Cantabrian Sea, southern Bay of Biscay). *Journal of Plankton Research*, *31*, 601–617.
Babin, M., & Stramski, D. (2004). Variations in the mass-specific absorption coefficient of mineral particles suspended in water. *Limnology and Oceanography*, *49*, 756–767.
Bakun, A., & Weeks, S.J. (2004). Greenhouse gas buildup, sardines, submarine eruptions and the possibility of abrupt degradation of intense marine upwelling ecosystems. *Ecology Letters*, *7*, 1015–1023.
Barton, A.D., Dutkiewicz, S., Flierl, G., Bragg, J., & Follows, M.J. (2010). Patterns of diversity in marine phytoplankton. *Science*, *327*, 1509–1511.
Beaugrand, G., Reid, P.C., Ibañez, F., Lindley, J.A., & Edwards, M. (2002). Reorganization of North Atlantic marine copepod biodiversity and climate. *Science*, *326*, 1692–1694.
Bélanger, S., Ehn, J.K., & Babin, M. (2007). Impact of sea ice on the retrieval of water-leaving reflectance, chlorophyll-a concentration and inherent optical properties from satellite ocean color data. *Remote Sensing of Environment*, *111*, 51–68.
Brandt, A., Gooday, A.J., Brandão, S.N., Brix, S., Brökeland, W., Cedhagen, T., et al. (2007). First insights into the biodiversity and biogeography of the Southern Ocean deep sea. *Nature*, *447*, 307–311.
Buddemeier, R.W., Smith, S.V., Swaney, D.P., Crossland, C.J., & Maxwell, B.A. (2008). Coastal typology: An integrative "neutral" technique for coastal zone characterization and analysis. *Estuarine, Coastal and Shelf Science*, *77*, 197–205.
Campbell, J.W. (1995). The log-normal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research*, *100*, 13237–13254.
Carr, M. -E., & Kearns, E.J. (2003). Production regimes in four eastern boundary current systems. *Deep Sea Research, II*, *50*, 3199–3221.
Castillo Atoche, A., Carrasco Alvarez, R., Ortegón Aguilar, J., & Vázquez Castillo, J. (2013). A new tool for intelligent parallel processing of radar/SAR remotely sensed imagery. *Mathematical Problems in Engineering*. http://dx.doi.org/10.1155/2013/405372.
Claustre, H., Morel, A., Hooker, S.B., Babin, M., Antoine, D., Oubelkheir, K., et al. (2002). Is desert dust making oligotrophic waters greener? *Geophysical Research Letters*, *29*, 1469. http://dx.doi.org/10.1029/2001GL14056.
D'Alimonte, D., Mélin, F., Zibordi, G., & Berthon, J. -F. (2003). Use of the novelty detection technique to identify the range of applicability of empirical ocean color algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, *41*, 2833–2843.
D'Ortenzio, F., & Ribera d'Alcalà, M. (2009). On the trophic regimes of the Mediterranean Sea: A satellite analysis. *Biogeosciences*, *6*, 139–148.

De Monte, S., Soccodato, A., Alvain, S., & d'Ovidio, F. (2013). Can we detect oceanic biodiversity hotspots from space? *The ISME Journal*. http://dx.doi.org/10.1038/ismej.2013.72.

Diaz, R.J., & Rosenberg, R. (2008). Spreading dead zones and consequences for marine ecosystems. *Science*, *321*, 926–929.

Dodson, S.I., Arnott, S.E., & Cottingham, K.L. (2000). The relationship in lake communities between primary productivity and species richness. *Ecology*, *81*, 2662–2679.

Doxaran, D., Cherukuru, N., & Lavender, S.J. (2006). Apparent and inherent optical properties of turbid estuarine waters: Measurements, empirical quantification relationships, and modeling. *Applied Optics*, *45*, 2310–2324.

Dürr, H.H., Laruelle, G.G., van Kempen, C.M., Slomp, C.P., Meybeck, M., & Middelkoop, H. (2011). Worldwide typology of nearshore coastal systems: Defining the estuarine filter of river inputs to the ocean. *Estuaries and Coasts*, *34*, 441–458.

Eiane, K., Aksnes, D.L., Bagøien, E., & Kaartvedt, S. (1999). Fish or jellies — A question of visibility? *Limnology and Oceanography*, *44*, 1352–1357.

Finkel, Z.V., Quigg, A., Raven, J.A., Reinfelder, J.R., Schofield, O.E., & Falkowski, P.G. (2006). Irradiance and the elemental stoichiometry of marine phytoplankton. *Limnology and Oceanography*, *51*, 2690–2701.

Fu, G., Baith, K.S., & McClain, C.R. (1998). SeaDAS: The SeaWiFS Data Analysis System. *Proceedings of the 4th Pacific Ocean Remote Sensing Conference, Qingdao, China, July 1998* (pp. 73–79).

Gallego, G., Cuevas, C., Mohedano, R., & García, N. (2013). On the Mahalanobis distance classification criterion for multidimensional normal distributions. *IEEE Transactions on Signal Processing*, *61*, 4387–4396.

Galloway, J.N., Townsend, A.R., Erisman, J.W., Bekunda, M., Cai, Z., Freney, J.R., et al. (2008). Transformation of the nitrogen cycle: Recent trends, questions, and potential solutions. *Science*, *320*, 889–892.

Gattuso, J. -P., Gentili, B., Duarte, C.M., Kleypas, J.A., Middelburg, J.J., & Antoine, D. (2006). Light availability in the coastal ocean: Impact on the distribution of benthic photosynthetic organisms and contribution to primary production. *Biogeosciences*, *3*, 489–513.

Gong, G. -C., Wen, Y. -H., Wang, B. -W., & Liu, G. -J. (2003). Seasonal variation of chlorophyll-a concetration, primary production and environmental conditions in the subtropical East China Sea. *Deep Sea Research, II*, *50*, 1216–1236.

Gordeev, V.V. (2006). Fluvial sediment flux in the Arctic Ocean. *Geomorphology*, *80*, 94–104.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Michelo, F., D'Agrosa, C., et al. (2008). A global map of human impact on marine ecosystems. *Science*, *319*, 948–952.

Huisman, J., van Oostveen, P., & Weissing, F.J. (1999). Species dynamics in phytoplankton blooms: Incomplete mixing and competition for light. *American Naturalist*, *154*, 46–68.

Irigoien, X., Huisman, J., & Harris, R.P. (2004). Global biodiversity patterns of marine phytoplankton and zooplankton. *Nature*, *429*, 863–867.

ISODATA (2014). http://www.exelisvis.com/docs/ISODATAClassification.html

Jensen, J.R. (1996). *Introductory digital image processing: A remote sensing perspective* (second ed.). Upper Saddle River, NJ: Prentice Hall.

Joint, I., Groom, S.B., Wollast, R., Chou, L., Tilstone, G.H., Figueiras, F.G., et al. (2002). The response of phytoplankton production to periodic upwelling and relaxation events at the Iberian shelf break: Estimates by the ^{14}C method and by satellite remote sensing. *Journal of Marine Systems*, *32*, 219–238.

Katsanevakis, S., Zenetos, A., Belchor, C., & Cardoso, A.C. (2013). Invading European seas: Assessing pathways of introduction of marine aliens. *Ocean and Coastal Management*, *76*, 64–74.

Kutser, T. (2004). Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. *Limnology and Oceanography*, *49*, 2179–2189.

Lantuit, H., Overduin, P.P., Couture, N., Wetterich, S., Aré, F., Atkinson, D., et al. (2012). The Arctic coastal dynamics database: A new classification scheme and statistics on Arctic permafrost coastlines. *Estuaries and Coasts*, *35*, 383–400.

Le Fouest, V., Babin, M., & Tremblay, J. -E. (2013). The fate of riverine nutrients on Arctic shelves. *Biogeosciences*, *10*, 3661–3677.

Lee, Z. -P., Shang, S., Hu, C., & Zibordi, G. (2014). Spectral interdependence of remote-sensing reflectance and its implications on the design of ocean color sensors. *Applied Optics*, *53*, 3301–3310.

Li, W.K.W. (2002). Macro-ecological patterns of phytoplankton in the northwest North Atlantic Ocean. *Nature*, *419*, 154–157.

Loisel, H., & Morel, A. (2001). Non-isotropy of the upward radiance field in typical coastal (Case 2) waters. *International Journal of Remote Sensing*, *22*, 275–295.

Longhurst, A.R. (2006). *Ecological geography of the sea*. Academic Press (560 pp.).

Lotze, H.K., Lenihan, H.S., Bourque, B.J., Bradbury, R.H., Cooke, R.G., May, M.C., et al. (2006). Depletion, degradation, and recovery potential of estuaries and coastal seas. *Science*, *312*, 1806–1809.

Lubac, B., & Loisel, H. (2007). Variability and classification of remote sensing reflectance spectra in the eastern English Channel and southern North Sea. *Remote Sensing of Environment*, *110*, 45–58.

Martin Traykovski, L.V., & Sosik, H.M. (2003). Feature-based classification of optical water types in the Northwest Atlantic based on satellite ocean color data. *Journal of Geophysical Research*, *108*, 3150. http://dx.doi.org/10.1029/2001JC001172.

Marty, J. -C., Chiavérini, J., Pizay, M. -D., & Avril, B. (2002). Seasonal and interannual variations in phytoplankton production at DYFAMED time-series station, northwestern Mediterranean Sea. *Deep Sea Research, II*, *49*, 2017–2030.

McClain, C.R., Cleave, M.L., Feldman, G.C., Gregg, W.W., Hooker, S.B., & Kuring, N. (1998). Science quality SeaWiFS data for global biosphere research. *Sea Technology*, *39*, 10–16.

McCullough, I.M., Loftin, C.S., & Sader, S.A. (2012). Combining lake and watershed characteristics with Landsat TM data for remote estimation of regional lake clarity. *Remote Sensing of Environment*, *123*, 109–115.

Mélin, F., Steinich, C., Gobron, N., Pinty, B., & Verstraete, M.M. (2002). Optimal merging of LAC and GAC data from SeaWiFS. *International Journal of Remote Sensing*, *23*, 801–807.

Mélin, F., Vantrepotte, V., Clerici, M., D'Alimonte, D., Zibordi, G., Berthon, J. -F., et al. (2011). Multi-sensor satellite time series of optical properties and chlorophyll-a concentration in the Adriatic Sea. *Progress in Oceanography*, *91*, 229–244.

Mélin, F., & Zibordi, G. (2007). Optically based technique for producing merged spectra of water-leaving radiances from ocean color remote sensing. *Applied Optics*, *46*, 3856–3869.

Memarsadeghi, N., Mount, D.M., Netanyahu, N.S., & Le Moigne, J. (2007). A fast implementation of the ISODATA clustering algorithm. *International Journal of Computational Geometry and Applications*, *17*, 71–103.

Molnar, J.L., Gamboa, R.L., Revenga, C., & Spalding, M.D. (2008). Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology And the Environment*, *6*. http://dx.doi.org/10.1890/070064.

Moore, T.S., Campbell, J.W., & Dowell, M.D. (2009). A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. *Remote Sensing of Environment*, *113*, 2424–2430.

Moore, T.S., Campbell, J.W., & Feng, H. (2001). A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, *39*, 1764–1776.

Moore, T.S., Dowell, M.D., Bradt, S., & Ruiz Verdu, R. (2014). An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. *Remote Sensing of Environment*, *143*, 97–111.

Moore, C.M., Suggett, D.J., Hickman, A.E., Kim, Y. -N., Tweddle, J.F., Sharples, J., et al. (2006). Phytoplankton photoacclimation and photoadaptation in response to environmental gradients in a shelf sea. *Limnology and Oceanography*, *51*, 936–949.

Morel, A., & Gentili, B. (2009). The dissolved yellow substance and the shades of blue in the Mediterranean Sea. *Biogeosciences*, *6*, 2625–2636.

Morel, A., & Maritorena, S. (2001). Bio-optical properties of oceanic waters: A reappraisal. *Journal of Geophysical Research*, *106*, 7163–7180.

Morel, A., & Prieur, L. (1977). Analysis of variations in ocean color. *Limnology and Oceanography*, *19*, 591–600.

Navarro, G., & Ruiz, J. (2006). Spatial and temporal variability of phytoplankton in the Gulf of Cádiz through remote sensing images. *Deep Sea Research, II*, *53*, 1241–1260.

Ning, X., Liu, Z., Cai, Y., Fang, M., & Chai, F. (1998). Physico-biological oceanographic remote sensing of the East China Sea: Satellite and in situ observations. *Journal of Geophysical Research*, *103*, 21623–21635.

Opsahl, S., Benner, R., & Amon, R.M.W. (1999). Major flux of terrigenous dissolved organic matter through the Arctic Ocean. *Limnology and Oceanography*, *44*, 2017–2023.

Press, W.H., Teukolsky, S.A., Vetterling, W. T., & Flannery, B.P. (1992). *Numerical recipes in C — The art of scientific computing* (2nd ed.). Cambridge University Press, 994.

Ragni, M., & Ribera d'Alcalà, M. (2004). Light as an information carrier underwater. *Journal of Plankton Research*, *26*, 433–443.

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., & Maury, O. (2013). Dynamic biogeochemical provinces in the global ocean. *Global Biogeochemical Cycles*, *27*, 1046–1058. http://dx.doi.org/10.1002/gbc.20089.

Rutherford, S., D'Hondt, S., & Prell, W. (1999). Environmental controls on the geographic distribution of zooplankton diversity. *Nature*, *400*, 749–753.

Sathyendranath, S., & Platt, T. (2007). Spectral effects in bio-optical control on the ocean system. *Oceanologia*, *49*, 5–39.

Sathyendranath, S., Prieur, L., & Morel, A. (1989). A three-component model of ocean colour and its application to remote sensing of phytoplankton pigments in coastal waters. *International Journal of Remote Sensing*, *10*, 1373–1394.

Shannon, C.E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423 (623–656).

Sherman, K., & Hempel, G. (2009). The UNEP large marine ecosystem report: A perspective on changing conditions in LMEs of the world's regional seas. *UNEP Regional Seas Report and Studies No. 182, United Nations Environment Program, Nairobi, Kenya*.

Shi, W., & Wang, M. (2012). Satellite views of the Bohai Sea, Yellow Sea, and East China Sea. *Progress in Oceanography*, *104*, 30–45.

Small, C., & Nicholls, R.J. (2003). A global analysis of human settlement in coastal zones. *Journal of Coastal Research*, *19*, 584–599.

Stewart, K.R., Lewison, R.L., Dunn, D.C., Bjorkland, R.H., Kelez, S., Halpin, P.N., et al. (2010). Characterizing fishing effort and spatial extent of coastal fisheries. *PLoS ONE*, *5*, e14451. http://dx.doi.org/10.1371/journal.pone.0014451.

Sydor, M., Gould, R.W., Arnone, R.A., Haltrin, V.I., & Goode, W. (2004). Uniqueness in remote sensing of the inherent optical properties of ocean water. *Applied Optics*, *43*, 2156–2162.

Thomas, A.C., & Strub, P.T. (2001). Cross-shelf phytoplankton pigment variability in the California Current. *Continental Shelf Research*, *21*, 1157–1190.

Ting, C.S., Rocap, G., King, J., & Chisholm, S.W. (2002). Cyanobacteria photosynthesis in the oceans: The origins and significance of divergent light-harvesting strategies. *Trends in Microbiology*, *10*, 134–142.

Tittensor, D.P., Mora, C., Jetz, W., Lotze, H.K., Ricard, D., Vanden Berghe, E., et al. (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, *466*, 1098–1101.

Uusitalo, L., Fleming-Lehtinen, V., Hällfors, H., Jaanus, A., Hällfors, S., & London, L. (2013). A novel approach for estimating phytoplankton biodiversity. *ICES Journal of Marine Science*, *70*, 408–417.

Vantrepotte, V., Loisel, H., Dessailly, D., & Mériaux, X. (2012). Optical classification of contrasted coastal waters. *Remote Sensing of Environment*, *123*, 306–323.

Vörösmarty, C.J., Meybeck, M., Fekete, B., Sharma, K., Green, P., & Syvitski, J.P.M. (2003). Anthropogenic sediment retention: Major global impact from registered river impoundments. *Global Planetary Change*, *39*, 169–190 (2003).

Voss, M., Dippner, J.W., Humborg, C., Hürdler, J., Korth, F., Neumann, T., et al. (2011). History and scenarios of future development of Baltic Sea eutrophication. *Estuarine, Coastal and Shelf Science*, *92*, 307–322.

Walsh, S.J., McCleary, A.L., Mena, C.F., Shao, Y., Tuttle, J.P., González, A., et al. (2008). QuickBird and Hyperion data analysis of an invasive plant species in the Galapagos Islands of Ecuador: Implications for control and land use management. *Remote Sensing of Environment*, *112*, 1927–1941.

Wang, H., Yang, Z., Saito, Y., Liu, J.P., & Sun, X. (2006). Interannual and seasonal variation of the Huanghe (Yellow River) water discharge over the past 50 years: Connections to impacts from ENSO events and dams. *Global Planetary Change*, *50*, 212–225.

Wissel, B., Boeing, W.J., & Ramcharan, C.W. (2003). Effects of water color on predation regimes and zooplankton assemblages in freshwater lakes. *Limnology and Oceanography*, *48*, 1965–1976.

Worm, B., Sandow, M., Oschlies, A., Lotze, H.K., & Myers, R.A. (2005). Global patterns of predator diversity in the open oceans. *Science*, *309*, 1365–1369.

Zettler, E.R., & Carter, J.C.H. (1986). Zooplankton community and species responses to a natural turbidity gradient in Lake Temiskaming, Ontario-Quebec. *Canadian Journal of Fisheries and Aquatic Sciences*, *43*, 665–673.

Zhou, F., Xuan, J., Huang, D., Liu, C., & Sun, J. (2013). The timing and magnitude of spring phytoplankton blooms and their relationship with physical forcing in the central Yellow Sea in 2009. *Deep Sea Research, II*, *97*, 4–15.

Zibordi, G., Berthon, J. -F., Mélin, F., & D'Alimonte, D. (2011). Cross-site consistent in-situ measurements for satellite ocean color applications: The BiOMaP radiometric dataset. *Remote Sensing of Environment*, *115*, 2104–2115.

Zibordi, G., Mélin, F., & Berthon, J. -F. (2006). A time series of above-water radiometric measurements for coastal water monitoring and remote sensing product validation. *IEEE Geoscience and Remote Sensing Letters*, *3*, 120–124.