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.