1. Introduction

It has long been recognized that transmission of light in sea water is essential to the productivity of the oceans. It provides the energy necessary for ocean currents, and the majority of marine life is supported by the thin layer of warm water near the ocean’s surface. Light plays a decisive role in the primary formation of biomass by oceanic chlorophyll-bearing marine plants, through the process of photosynthesis, which is the basis of the entire marine food chain. Light transmission is therefore a key factor in the ecology of the upper ocean and biogeochemical cycling, since it has a strong influence on the dynamics of the chemical compounds. In addition, variability of light in the sea is strongly influenced by the distribution of the components in the water column, which varies across horizontal and vertical space and time scales.

The subdiscipline of ocean optics, which concerns the assessment of the propagation of light through the oceanic water column and surface, has become fundamental for understanding water dynamics and composition. Over the last decades, there has been a noticeable increase in the number of coastal and open-ocean studies using in situ and remote sensing optical measurements. Recent advances in monitoring surface and underwater optical properties have been specially related to the progressive shift from using multispectral to high spectral resolution (hyperspectral) acquisition systems. Hyperspectral technology has opened the possibility for optical oceanographers to more accurately characterize complex oceanic environments (Chang et al., 2004). A broad range of hyperspectral sensors, covering from several hundreds to thousands of contiguous spectral bands, have recently been developed for different monitoring applications. The use of high sample rate hyperspectral systems may provide the potential for mapping phytoplankton functional types, including the detection of harmful algal blooms (Nair et al., 2008; Craig et al., 2006).
The higher spectral resolution provides the opportunity to better perform spectral shape analysis of oceanic optical measurements. Techniques such as derivative spectroscopy, which serves to enhance subtle features in spectra, have been extensively used to assess information regarding optically significant water constituents. Derivative of hyperspectral data can yield more information than traditional analysis based on ratios of discrete spectral bands (multispectral approaches). However, when hyperspectral measurements are to be used for further analysis, the uncertainties of the measurement system must be taken into account. Appropriate calibration strategies must be followed, since the accuracy of the measurements relies on the associated uncertainties of the measurement system. A well-calibrated spectroradiometer includes a characterization process, describing the instrument behavior in terms of the spectral sensitivity, the signal-to-noise ratio, the dark current, the wavelength calibration, the nonlinearity, the temperature dependence of the measurements and the spectral scattering, or what is called the spectral stray-light (Brown et al., 2006). The stray-light of a spectroradiometer is described as the unwanted background radiation, scattered due to imperfections in the dispersive device and other internal optical elements.

In ocean optics applications, special attention must be paid to corrections for the inherent distortions of the hyperspectral sensors (e.g. noise, spectral stray-light), because the errors in the measured radiances distributions may be potentially significant and lead to inaccurate retrievals of water properties. This issue becomes even more important when commonly derivative spectroscopy is used to explore subtle features in shape of hyperspectral data (Torrecilla et al., 2008a). Derivative analysis is notoriously sensitive to noise and smoothing techniques must be used to overcome this problem. Therefore, in order to make an optimal application of the derivative analysis, it is worth noting that a suitable selection of the smoothing and derivative parameters (filter size and band separation) must be done according to the resolution of each type of hyperspectral data (Torrecilla et al., 2007). An effort must be made to determine the best trade-off between denoising and the ability to resolve spectral features of interest.

There are two major goals in this chapter. The first one is to offer a brief review of relevant advances that have been achieved using hyperspectral technology in oceanography and a description of some important issues that must be considered when derivative spectroscopy is applied to data acquired by hyperspectral sensors. The second goal is to provide some results demonstrating the feasibility of applying derivative spectroscopy to hyperspectral measurements of remote-sensing reflectance in open-ocean environments, in particular to identify phytoplankton pigment assemblages. In this part, a simulation-based framework is proposed to address this issue and special attention is given to the role of involved parameters when derivative analysis of hyperspectral data is performed. For instance, the effect of spectral sensitivity of each high spectral resolution instrument considered on the results of the derivative-based approach is also discussed.

2. Ocean Optics: an overview

In order to understand light propagation in the ocean, which is an optically complex environment, the radiative transfer theory was defined as the theoretical framework (Kirk, 1994; Mobley, 1994). Optical properties of sea water are classified into inherent and apparent...
optical properties (IOPs, AOPs) and are related via the radiative transfer equation (RTE). Most of the quantities are shown in Figure 1 and have wavelength dependence.

Fig. 1. Schematic diagram illustrating the relationships among the ocean optics quantities, based on Figure 3.27 of Mobley, 1994.

Spectral absorption and scattering coefficients are inherent optical properties (IOPs), i.e. they are only affected by the constituents of the aquatic medium and are independent of the ambient illumination conditions. IOPs are additive and their variation is generally attributed to variability in four constituents of the aquatic medium in natural oceanic waters: pure sea water, phytoplankton, detritus and gelbstoff or colored dissolved organic matter (CDOM). In contrast, the apparent optical properties (AOPs), such as spectral radiance and irradiance, are not additive and depend both on the medium and on the directional structure of the ambient light field. Several types of models have been reported to predict AOPs given field measurements of IOPs and environmental conditions (“direct models”) or to estimate IOPs given experimental measurements of AOPs (“inverse models”) (IOCCG, 2000).

In situ and remote sensed optical observations of IOPs and AOPs are two complementary oceanographic applications that offer the potential of being direct proxies for important biological and biogeochemical variables in the ocean (Chang et al., 2006). Remote sensing provides global-scale optical data, whereas in situ measurements are useful for calibration of remote sensing measurements and for obtaining higher temporal resolution data. Both types of observations of the ocean are important for providing relevant information regarding the relative concentrations of optically significant constituents in the water column. Several studies have traditionally focused on the development of bio-optical algorithms linking measurable optical properties to the primary pigment in phytoplankton, chlorophyll-a, a proxy for the phytoplankton biomass (Bricaud et al., 1998; Reynolds et al., 2001; O’Reilly et al., 2000), and on the detection of phenomena such as marine harmful algal blooms (HABs) (Cullen et al., 1997). However, other approaches have been more recently
applied to estimate the surface concentration of particulate organic carbon (POC) from optical measurements (Stramski et al., 2008), to derive the vertical distribution of phytoplankton communities in the open-ocean based on the near-surface chlorophyll-\(a\) content (Uitz et al., 2006), and to map phytoplankton functional types (PFTs) from ocean-color data (Nair et al., 2008), which serve as a contribution to the mapping of biodiversity in marine phytoplankton on the global scale.

3. Hyperspectral Data Acquisition Systems in Oceanography

Technological advances and innovations in instrumentation have given rise during the last few decades to a broad range of optical sensors for different monitoring applications in oceanography, including active and passive systems. Passive radiometric systems have been widely used to measure properties such as radiance and irradiance distributions, and solar-induced fluorescence. They can operate only with sunlight and are restricted to measurements within the euphotic depth (i.e. to a depth where light intensity falls to 1% of that at the surface). On the other hand, active radiometric systems use an internal light source to measure other optical properties such as absorption, scattering and fluorescence. These systems are able to provide optical measurements at deeper depths and at wavelengths that have reduced penetration in the ocean (e.g. the UV portion of the electromagnetic spectrum). However, their use involves a higher power consumption.

The rapid maturing of optical instrumentation has led to the most significant recent advance which is the development of hyperspectral sensors. The ability to measure surface and underwater light fields at hundreds of narrow and closely spaced wavelength bands, with a resolution better than 10 nm, has become one of the most powerful and fastest growing technologies in the field of ocean optics. Relative to multispectral measurements, hyperspectral sensors potentially enable the extraction of more detailed spectral information than is possible with any other type of in situ and remote sensing optical data. Figure 2 shows examples of phytoplankton absorption (IOP) and remote-sensing reflectance spectra (AOP), gathered by multi- and hyperspectral systems. Hyperspectral patterns of absorption or reflectance across wavelengths provide more information about spectral singularities.

The advantage of a hyperspectral over a multispectral inversion for phytoplankton species identification stems from the fact that more accurate spectral information, such as spectral features related to characteristic phytoplankton pigment absorption peaks, is resolved. In the past, analyses of single band-ratios obtained in discrete multispectral bands were employed to resolve a variety of water column properties (e.g. chlorophyll-\(a\) concentration (O’Reilly et al., 2000), the presence of specific species and other water constituents (Sathyendranath et al., 2004), etc.). However, the advent of hyperspectral sensors provided ample spectral information and facilitated the identification of water column constituents, such as phytoplankton species, which are spectrally unique. For instance, Craig et al. (2006) assessed the feasibility of remote detection and monitoring of the toxic dinoflagellate, Karenia brevis, applying two numerical methods to in situ hyperspectral measurements of remote-sensing reflectance. Bracher et al. (2008) recently adapted the differential optical absorption spectroscopy (DOAS) technique for retrieving the absorption and biomass of two major phytoplankton groups of cyanobacteria and diatoms from high spectrally resolved
data of the satellite sensor SCIAMACHY (Scanning Imaging Absorption Spectrometer for Atmospheric Chartography). Lohrenz et al. (2008) directly explored the hyperspectral patterns of remote-sensing reflectance to better characterize water mass properties in coastal areas, and Gagnon et al. (2008) performed sea-bed mapping of benthic assemblages in optically complex shallow waters.

Fig. 2. (Left) Hyperspectral (solid continuous line, 1 nm resolution) and multispectral (closed circles) phytoplankton absorption spectra. (Right) Hyperspectral (solid line, 1 nm resolution) and multispectral (closed circles) remote-sensing reflectance spectra. (Absorption and multispectral remote-sensing reflectance data were provided by Dariusz Stramski and Rick Reynolds, Scripps Institution of Oceanography, from field measurements during the ANT-XXIII/1 expedition on R/V Polarstern in the eastern Atlantic in 2005 (Stramski et al., 2008)).

3.1 Hyperspectral Sensors
Hyperspectral sensors generally use diffraction gratings or linear variable optical filters as a dispersive element to separate light into specific wave bands centered on desired wavelengths, and different scanning mechanisms to generate one- or two-dimensional high-resolution spectra or images. Mechanical-scanning spectrometers have been traditionally employed to obtain a spectrum by using a single photo-detector and rotating the dispersive element over a given specific spectral region after passing through the sample. However, more recent spectrometers use a fixed grating and multi-element array detectors that allow an entire spectrum to be acquired over some finite spectral region simultaneously.

Array spectrometers are being widely used as a tool for rapid measurements of spectral distributions in oceanographic applications (Figure 3) in which acquisition speed is an important issue. Spectra can be measured using three types of arrays: charge-coupled device (CCD), photodiode or complementary metal-oxide semiconductor (CMOS). High performance is available in all technologies today when they are designed properly, and each have their own strengths and weaknesses. For instance, CCD arrays show a slightly higher dynamic range (sensitivity) and a lower system noise than CMOS ones, whereas CMOS ones offer more integration, smaller system sizes and higher speeds. The proper
selection of the type of array depends on many parameters (e.g. pixel dimensions, sensitivity, spectral range coverage, dynamic range, saturation exposure, integration time) and on the specific application. Other important advantages of array spectrometers are the non moving parts, the robustness and the low production costs. A key challenge has also been to construct small size array spectrometers without compromising performance. However, they involve several drawbacks compared with mechanical-scanning spectrometers, such as the fixed wavelength resolution, the lower sensitivity and the higher stray-light radiation (described in more detail below), due to the lack of an output slit and the integral illumination over the full wavelength range.

New technologies and the miniaturization of electro-optical components have permitted the development of accurate, low-cost and energy-efficient hyperspectral sensors designed to measure oceanic high spectral resolution IOPs and AOPs (absorption, total upwelling or downwelling radiance and irradiances, etc.). In the design and development of an instrument for hyperspectral measurements in the ocean, it is necessary to consider several optimal characteristics required of its detection system. The wavelength coverage should be broad, ideally from 350 to 750 nm. A low-light-level detector should be selected with a large responsivity should be high over the whole spectral region covered, especially in the blue and red spectral portions, and the dark signal should be low. Furthermore, in order to minimize a variety of external perturbations during measurement (e.g. changing sky conditions, surface wave noise producing high-frequency fluctuations in irradiance, ship shadow effect), the scan time of the system should be fast. The fastest devices are multi-element array spectrometers with a fixed grating system.

3.2 Oceanographic Platforms for Hyperspectral Sensors
Currently available miniature spectrometers based on linear array detectors represent best possible compromises between size, cost, resolution, range and performance. As a result, their use is no longer limited to bench-top applications but they are being deployed on fully fledged sea-going observational platforms. New interdisciplinary ocean observing platforms incorporating hyperspectral sensors are being deployed (Dickey et al., 2006). Various satellite-based hyperspectral imagers are currently available: NASA's Hyperion sensor on
the EO-1 satellite, ESA's CHRIS sensor on PROBA satellite or the U.S. Air Force Research Lab's FTHSI sensor on the MightySat II satellite. In addition, examples of existing hyperspectral airborne sensors providing high spatial resolution images of specific areas are the Ocean Portable Hyperspectral Imager for Low-Light Spectroscopy (PHILLS, Davis et al., 2002), the Compact Airborne Spectrographic Imager (CASI), the Airborne Visible Infrared Imaging Spectrometer (AVIRIS), etc. Other in situ ocean observing platforms for which hyperspectral sensors may be suitable are ships and moored buoys (Kuwahara et al., 2007). These platforms offer continuous measurements at a high spatial and temporal resolution delivering data even under cloudy conditions, thereby complementing discrete observations by satellites and providing the data essential for calibration and validation purposes. Furthermore, Lagrangian platforms that follow a particular water mass (e.g. floats and drifters) can provide hyperspectral measurements in oceanic regions not normally accessible by satellites or oceanographic research vessels. The goal of these research initiatives is to determine the material composition of the water mass under investigation through the accurate analysis of simultaneous optical and hydrographic measurements.

The number of newly emerging ocean observing platforms capable of incorporating hyperspectral instrumentation is continuously growing. For instance, new autonomous underwater vehicles (AUVs) have been developed recently with integrated equipment for adaptive sampling and the ability to perform a wide-range of preprogrammed monitoring surveys over extended periods of time (Perry & Rudnick, 2003). Another member of this family of instruments are the special AUVs called gliders that propel themselves through the water by changing their buoyancy with a minimal expenditure of energy. The glider provides a wide spatial coverage measuring - among others - the optical properties during periods of three or four weeks (Woods, 2009). Another approach to monitor the variability of optical properties in the ocean is the use of profiling systems which can, for instance, be part of some of the emerging cabled observatories. Instrument platforms such as the Vertical Profiler System (VPS), part of Neptune Canada (Neptune, 2009), the world’s first regional-scale underwater ocean observatory plugged directly into the Internet, offer great power and bandwidth for gathering large quantities of hyperspectral measurements.

The amount of hyperspectral data sets available collected using all these types of platforms, covering a wide range of temporal and spatial scales, will increase in the near future. New interdisciplinary research initiatives are being successfully carried out such as the Hyperspectral Coastal Ocean Dynamics Experiment (HyCODE, 2009) conceived to exploit the new capabilities of hyperspectral ocean color sensors. Within this framework, several collaborative short-term and long-term field experiments, involving the simultaneous use of hyperspectral devices deployed in situ and remotely, are being conducted for calibrating, groundtruthing and relating subsurface optical properties to remote sensing ocean color measurements. These investigations are essential to develop and validate optical models and to further our understanding of the processes that contribute to the temporal and spatial variability of IOPs and AOPs in the ocean. One of the most important challenges of HyCODE experiment is the standardization of the quality control protocols, including the definition of procedures to perform detailed analysis of the reliability of measurements and an evaluation of the impact of instrumentations inaccuracies.
4. Hyperspectral Data Pre- and Post-Processing Methods in Oceanography

4.1 Data Quality Assurance

When hyperspectral measurements can be performed and are to be used for further analysis, the uncertainties of the measurement system must always be taken into account. Appropriate calibration and pre-processing strategies must be undertaken, since the accuracy of the measurements relies on the associated uncertainties of the measurement system. The clearest example in optical oceanography is perhaps the in situ radiometric measurements carried out during several field campaigns, which are still currently essential to provide a proper calibration and validation of remote sensing measurements collected by ocean color satellites and airborne platforms (Zibordi et al., 2001; McClain et al., 2004; Clark et al., 2003).

A well-calibrated high spectral resolution array-based spectrometer, for both in situ and remote sensing applications, should include likewise its characterization and an evaluation of all meaningful sources of uncertainty (Lewis, 2008; Voss et al., 2008). It is therefore always necessary to describe the instrument's behavior in terms of responsivity, signal-to-noise ratio (SNR), dark current, nonlinearity, temperature dependence of measurements and spectral scattering, or what is called the spectral stray-light of an instrument.

Changes in the thermal response of these systems are due to the silicon used for the arrays of detectors. For instance, within the framework of the SORTIE (Spectral Ocean Radiance Transfer Investigation Experiment) project, corrections to in situ hyperspectral radiometric measurements were performed and demonstrated to be very repeatable (McLean, 2008). The thermal characterization was carried out using shutters and installing a thermistor on the diode array. The spectral scattering or stray-light radiation of an array-based spectrometer is described as the unwanted background radiation that has been scattered due to imperfections in the fixed dispersive element and other optical elements, such as higher-order diffraction gratings, surfaces and internal baffles. The characterization of an instrument's response for measurement errors arising from its spectral stray-light can be performed using a set of monochromatic spectral line sources covering the entire instrument's operational spectral range. This method is based on computing the ratio of the spectral stray-light signal to the total signal within the bandpass of an array-based spectrometer (Brown et al., 2006). Figure 4 shows an example of the effectiveness of this spectral stray-light correction method applied to the hyperspectral CMOS-array spectrometer from Figure 3 (Torrecilla et al., 2008a). The correction method has been proved to be effective using a calibration broadband light source and a green absorption bandpass filter. The stray-light signals outside the filter's bandpass region are clearly reduced by more than two orders of magnitude, to a level of $10^{-4}$. 

Fig. 4. Spectral stray-light correction applied to the miniature hyperspectral CMOS-array spectrometer in Figure 3, using a broadband source with a bandpass filter. Normalized measured and corrected signals. The y-axis is a logarithmic scale.

Fig. 5. (Left) Spectral stray-light correction applied to field underwater light measurements at two depths, acquired by a miniature hyperspectral CCD-array spectrometer. (Right) Difference in percentage between measured and corrected signals at each depth.
4. Hyperspectral Data Pre- and Post-Processing Methods in Oceanography

4.1 Data Quality Assurance
When hyperspectral measurements can be performed and are to be used for further analysis, the uncertainties of the measurement system must always be taken into account. Appropriate calibration and pre-processing strategies must be undertaken, since the accuracy of the measurements relies on the associated uncertainties of the measurement system. The clearest example in optical oceanography is perhaps the in situ radiometric measurements carried out during several field campaigns, which are still currently essential to provide a proper calibration and validation of remote sensing measurements collected by ocean color satellites and airborne platforms (Zibordi et al., 2001; McClain et al., 2004; Clark et al., 2003).

A well-calibrated high spectral resolution array-based spectrometer, for both in situ and remote sensing applications, should include likewise its characterization and an evaluation of all meaningful sources of uncertainty (Lewis, 2008; Voss et al., 2008). It is therefore always necessary to describe the instrument’s behavior in terms of responsivity, signal-to-noise ratio (SNR), dark current, nonlinearity, temperature dependence of measurements and spectral scattering, or what is called the spectral stray-light of an instrument. Changes in the thermal response of these systems are due to the silicon used for the arrays of detectors. For instance, within the framework of the SORTIE (Spectral Ocean Radiance Transfer Investigation Experiment) project, corrections to in situ hyperspectral radiometric measurements were performed and demonstrated to be very repeatable (McLean, 2008). The thermal characterization was carried out using shutters and installing a thermistor on the diode array. The spectral scattering or stray-light radiation of an array-based spectrometer is described as the unwanted background radiation that has been scattered due to imperfections in the fixed dispersive element and other optical elements, such as higher-order diffraction gratings, surfaces and internal baffles. The characterization of an instrument’s response for measurement errors arising from its spectral stray-light can be performed using a set of monochromatic spectral line sources covering the entire instrument’s operational spectral range. This method is based on computing the ratio of the spectral stray-light signal to the total signal within the bandpass of an array-based spectrometer (Brown et al., 2006). Figure 4 shows an example of the effectiveness of this spectral stray-light correction method applied to the hyperspectral CMOS-array spectrometer from Figure 3 (Torrecilla et al., 2008a). The correction method has been proved to be effective using a calibration broadband light source and a green absorption bandpass filter. The stray-light signals outside the filter’s bandpass region are clearly reduced by more than two orders of magnitude, to a level of $10^{-4}$.

Fig. 4. Spectral stray-light correction applied to the miniature hyperspectral CMOS-array spectrometer in Figure 3, using a broadband source with a bandpass filter. Normalized measured and corrected signals. The y-axis is a logarithmic scale.

A data pre-processing and corrections for the spectral stray-light must be given special attention for underwater optical measurement purposes, in which very weak signals are collected and the errors in the measured radiometric distributions may be potentially significant, leading to inaccurate retrievals of water properties. For instance, Figure 5 (left panel) depicts measured and spectral stray-light corrected signals corresponding to field underwater light measurements gathered by a miniature hyperspectral CCD-array spectrometer at two different depths at a test site in the Alfacs Bay (Ebre Delta, NW Mediterranean). It is worth noting that the difference in percentage (Figure 5, right panel) between measured and spectral stray-light-corrected data at each depth has a spectral dependency (i.e. it is not the same for each spectral band). Therefore, the amplitude and shape of signals is modified when the spectral stray-light correction is applied. This is an issue that becomes very important when spectral shape analysis techniques such as the central topic of this chapter derivative spectroscopy, which is described below are used to explore subtle features in hyperspectral data.

Fig. 5. (Left) Spectral stray-light correction applied to field underwater light measurements at two depths, acquired by a miniature hyperspectral CCD-array spectrometer. (Right) Difference in percentage between measured and corrected signals at each depth.
In addition to the above-mentioned corrections, new spectrometric devices and deployment techniques are now being developed, in order to optimize spectrometer's throughput and accuracy. Some spectrometers are designed with a two-dimensional area detector. They offer a significant improvement in the signal-to-noise ratio by averaging a vertical row of pixels to obtain each value of the whole spectrum at each wave band. Another solution is a new deployment technique called the multicast technique, which consists in performing and processing several casts in order to achieve the highest accuracy when underwater radiometric measurements are collected by hyperspectral devices (McLean, 2008). These are critical issues that should be tackled in addition to the radiometric and wavelength calibrations that must be carried out periodically using standard and highly characterized lamps (e.g. at the National Institute of Standards and Technology [NIST]).

4.2 Derivative Analysis

The emerging hyperspectral technology in remote sensing and in situ optical oceanography leads to a need for ongoing evaluation and improvement of hyperspectral processing methods. The higher spectral resolution provides the opportunity to develop and evaluate advanced methods of spectral shape analysis, such as derivative spectroscopy, that better distinguish subtle features in spectra and may be critical for discriminating optically significant water constituents.

Derivative spectroscopy has been commonly used in the analysis of hyperspectral data using different computation algorithms (Tsai & Philpot, 1998; Ruffin et al., 2008). The process of estimating derivative spectra can be addressed using a finite divided difference algorithm, named “finite approximation”, which consists in computing the changes in curvature of a given spectrum over a sampling interval ($\Delta \lambda$) or band separation ($BS$), defined as $\Delta \lambda = \lambda_j - \lambda_i$, where $\lambda_j > \lambda_i$. The first and the $n$th derivative are obtained using Eqs. 1 and 2, respectively:

$$\frac{ds}{d\lambda} \approx \frac{s(\lambda_i) - s(\lambda_j)}{\Delta \lambda} \quad (1)$$

$$\frac{d^n s}{d\lambda^n} \approx \frac{d}{d\lambda} \left( \frac{d^{(n-1)} s}{d\lambda^{(n-1)}} \right) \quad (2)$$

Derivative analysis can be applied to hyperspectral measurements of both inherent and apparent oceanographic optical properties (e.g. absorption, irradiance, remote-sensing reflectance). It is a useful tool for enhancing spectral features, which are often relevant, for instance, because they can be related to absorption bands of pigments present in the considered water samples (Figure 6, left panels). Derivatives of hyperspectral data are less affected by possible spectral fluctuations of skylight conditions. However, successful extraction of spectral details of interest through derivative spectroscopy depends on the band separation chosen. The importance of selecting a suitable band separation ($\Delta \lambda = BS$) in each case stems from the fact that spectral data features of interest with a smaller scale than the band separation will not be preserved in the derivative results. Figure 6 (right panel)
shows different derivatives of the spectrum shown in the left panel, computed according to different finite band resolutions or band separations ($BS$).

Fig. 6. (Left) Original spectrum and its second derivative of a water sample containing a unique phytoplankton algal culture ($Alexandrium minutum$), acquired by a miniature hyperspectral CCD-array spectrometer. (Right) Second derivatives computed for several values of band separation ($BS$), each of them leading to spectral features at different scales.

Noise level in hyperspectral data can be considerable, as the small amount of energy gathered by the narrow bandwidth may be exceeded by the intrinsic sensor's noise. In order to make an optimal application of derivative analysis, which is a technique clearly sensitive to noise, smoothing techniques must be applied previously to derivative computation of hyperspectral data (Vaiphasa, 2006). A number of smoothing algorithms have been developed within the last few decades (e.g. Savitzky-Golay, Kawata-Minami or mean-filter smoothing). In all approaches the smoothing level applied depends on the size of the filter window ($WS$). It is therefore worth noting that an appropriate selection of the smoothing and derivative parameters (filter size and band separation) must be done according to the resolution of each type of hyperspectral data (Torrecilla et al., 2007). An important effort must be made to determine the best compromise between denoising and the ability to resolve fine spectral details of interest.

The advantages offered by hyperspectral measurements of remote-sensing reflectance ($R_{rs}(\lambda)$) in combination with derivative spectroscopy have been recently exploited for various purposes in optical oceanography. For example, Louchard et al. (2002) assessed qualitative and quantitative information regarding major sediment pigments of benthic substrates from derivatives of shallow marine hyperspectral $R_{rs}(\lambda)$. In addition, a bathymetric algorithm was suggested through the analysis of the magnitude of some derivative peaks. Craig et al. (2006) detected a toxic algal bloom from the analysis of the fourth derivative of phytoplankton absorption spectra, estimated from in situ hyperspectral measurements of $R_{rs}(\lambda)$ using a quasi-analytical inversion algorithm. The monitoring of these harmful algal blooms is possible because some accessory pigments are unique to individual phytoplankton taxa and can be better differentiated in hyperspectral absorption spectra than in multispectral spectra with a limited number of wavelengths. The advantage of a hyperpectral over a multispectral inversion for identification of other algal blooms was also confirmed by Lubac et al. (2008). In this case, the inversion and quantitative assessment are directly based on the analysis of the position of the maxima and minima of the second derivative of $R_{rs}(\lambda)$. 

$$\frac{\Delta}{d\lambda}$$
4.3 A model-based approach to evaluating the efficacy of processing methods

Most of the studies dedicated to derivative analysis of hyperspectral remote-sensing reflectance have traditionally focused on the use of ratios of discrete values of derivative spectra to characterize or classify oceanic environments. However, more advanced and contemporary methods are beginning to take advantage of the information contained in the whole derivative spectrum. For example, Filippi (2007) proposes a combination of derivative spectroscopy and artificial neural network (ANN) algorithms. The derivative-neural approach has been proven to be effective for providing bathymetry, bottom type and constituent concentration estimations from remote-sensing reflectance \( (R_{rs}(\lambda)) \) measurements.

In order to test the effectiveness of such a methodology, or the one described in this chapter, it would be desirable to have a large number of observed oceanographic hyperspectral data sets, covering a wide range of environmental conditions. However, because valid hyperspectral data sets are still difficult to obtain and often unavailable, oceanic radiative transfer (RT) models are used. Hydrolight is an example of a radiative transfer numerical model (Mobley, 1994; Mobley & Sundman, 2008), which computes radiance distributions and derived quantities (e.g. \( R_{rs}(\lambda) \)) given water column IOPs and other oceanographic environmental conditions. The Hydrolight code employs mathematically sophisticated invariant imbedding techniques to solve the radiative transfer equation (RTE, Figure 1) and offers the possibility of performing numerical simulations in controlled environments (Albert & Mobley, 2003; Kempeneers et al., 2005). For instance, Figure 7 depicts several Hydrolight-generated hyperspectral \( R_{rs}(\lambda) \) spectra, corresponding to different open-ocean scenarios, each dominated by a single phytoplankton group.

![Fig. 7. Results of Hydrolight radiative transfer simulations, showing four remote-sensing reflectances \( R_{rs}(\lambda) \) (1 nm resolution) obtained under different composition conditions. Labels indicate each dominant phytoplankton group considered in each case.](image)

One of the logical steps for improving the Hydrolight-based approaches, when simulated hyperspectral \( R_{rs}(\lambda) \) are to be used as a basis for further validation of some processing techniques, is to conduct sensor experiments. With the goal of exploring the potential of any processing technique, the response of the sensor with which the \( R_{rs}(\lambda) \) acquired would be
acquired must be taken into account. More accurate and realistic retrievals will be carried out from hyperspectral oceanographic data if, as has been stated as a priority in Section 4.1, the effect of the sensor in terms of noise, sensitivity, spectral resolution, stray-light, etc. is included in the simulation-based approach. This is one of the main issues that will be discussed in the next section.

5. Experimental Results

5.1 Experimental Design
In the following study, we investigate the potential offered by hyperspectral sensors and derivative spectroscopy to identify phytoplankton assemblages from hyperspectral measurements of remote-sensing reflectance \( R_{\text{rs}}(\lambda) \) in open-ocean waters. A simulation-based framework was used to achieve this goal, which includes the use of the Hydrolight-Ecolight version 5 radiative transfer model (Mobley & Sundman, 2008). The analysis methodology of this research is presented schematically in Figure 8.

![Flowchart showing the methodology followed for testing the automatic identification of phytoplankton assemblages from hyperspectral \( R_{\text{rs}}(\lambda) \) and derivative spectroscopy.](image)

The generation of a set of Hydrolight-simulated hyperspectral \( R_{\text{rs}}(\lambda) \) spectra (1 nm resolution) takes into account the constituents present in the water column and their inherent optical properties (absorption, scattering, etc.). The sensor’s behavior can also be included in the \( R_{\text{rs}}(\lambda) \) computation (see dashed box in Figure 8). The next step defined in the diagram is the computation of the second derivative of normalized \( R_{\text{rs}}(\lambda) \) spectra (i.e. \( \frac{d^2}{d\lambda^2} R_{\text{rs}}(\lambda) \)) normalized by the \( R_{\text{rs}}(\lambda) \) obtained at \( \lambda=555\)nm. A normalized \( R_{\text{rs}}(\lambda) \) is used as an input in order to emphasize the shape singularities of each spectrum before derivative spectroscopy. Previous to derivative analysis, a mean filter smoothing type is also applied to normalized \( R_{\text{rs}}(\lambda) \) data. It consists of a simple average of points within the chosen filter window. As stated in Section 4.2, the efficient way to accomplish the smoothing and derivative is to carefully adapt the filter size and the sampling interval (BS) to better match the scale of the spectral features of interest in each case.

Finally, the goal of validating the potential of hyperspectral \( R_{\text{rs}}(\lambda) \) measurements for identifying phytoplankton assemblages is addressed through the comparison of second
derivative spectra instead of discrete values of $R_{rs}(\lambda)$ spectra themselves. To make such a comparison, an approach based on hierarchical cluster analysis (HCA) is used (Jain et al., 1999). HCA is a common methodology consisting in creating a hierarchical cluster tree to partition a data set into subsets (clusters) using a single linkage algorithm. The linkage algorithm is based on a previously calculated pairwise distance between observations (i.e. each second derivative of a normalized $R_{rs}(\lambda)$ spectrum). The selected distance measure determines how the similarity of two spectra is calculated. In this case, one minus the cosine of the included angle between two vectors was used as a distance measure (cosine distance).

As a linkage algorithm, the shortest distance between vectors, also called the nearest neighbor (single linkage), was selected. The traditional representation of this hierarchical tree is a dendrogram, with individual elements at one end and a single cluster containing every element at the other. Note that the smaller the cosine distance between two observations, the more similar are the features of the two compared derivatives of normalized $R_{rs}(\lambda)$ spectra. Therefore, spectra corresponding to observations with a similar phytoplankton composition are expected to appear closer in the dendrogram than those having a very different phytoplankton composition. The feasibility of applying derivative spectroscopy to hyperspectral measurements of $R_{rs}(\lambda)$ to identify phytoplankton assemblages will be assessed by analyzing how close spectra belonging to similar phytoplankton assemblages appear in the computed cluster tree (or dendrogram).

5.2 Automatic Identification of Phytoplankton Assemblages

The results shown in this subsection provide the first detailed demonstration of the advantages and limitations of integrating hyperspectral $R_{rs}(\lambda)$ data and derivative spectroscopy for the automatic identification of phytoplankton assemblages in open-ocean waters. A hyperspectral $R_{rs}(\lambda)$ data set, covering a range of environmental conditions, was created using absorption and scattering properties of six different phytoplankton groups (Figure 9, left panel) (Kim and Philpot, 2006).

![Fig. 9. (Left) Specific absorption spectra of six phytoplankton groups. (Right) Simulated hyperspectral $R_{rs}(\lambda)$ spectra (1 nm resolution) corresponding to different dominating phytoplankton groups at different concentrations (0.01, 0.03, 0.05, 0.07 and 0.09 mg/m$^3$).]
A total of thirty hyperspectral $R_n(\lambda)$ spectra were simulated, each of them dominated by a single phytoplankton group along the water column, which was assumed to be homogeneous. The spectra were generated from combining six different dominating phytoplankton groups and five different concentration values (Figure 9, right panel) corresponding to typical concentrations encountered under non-bloom sea conditions. It must be noted that $R_n(\lambda)$ spectra display great variability in both magnitude and spectral shape, in correspondence with the variable phytoplankton composition and concentration rates.

Spectral derivatives and clustering techniques (HCA) were used independently to identify and group similar phytoplankton assemblages from hyperspectral $R_n(\lambda)$ spectra. Figure 10 shows the results of cluster analysis applied to the raw hyperspectral normalized $R_n(\lambda)$ spectra (top panel) and to the second derivative of hyperspectral normalized $R_n(\lambda)$ spectra (central panel). Each simulated $R_n(\lambda)$ is identified with a specific label, consisting of the name of the dominating phytoplankton group and the concentration value. For instance, if diatoms are the dominating phytoplankton group, with a concentration of 0.05 mg/m$^3$ along the homogeneous water column, the label identifying that case will be Diat_0.05. It is worth noting that when the raw normalized $R_n(\lambda)$ is used for cluster analysis (Figure 10, top panel), the only group of $R_n(\lambda)$ spectra which is clustered satisfactorily is the one corresponding to the Prasinophyceae phytoplankton group (see labels in cyan color). The remaining cases, from different phytoplankton groups, are mixed. However, when second derivatives of normalized $R_n(\lambda)$ are used, the majority of phytoplankton groups at different concentration rates are identified and connected by closer dendrites in the dendrogram. Therefore, derivatives of spectra from the same phytoplankton group (and color) are located closer in the dendrogram. The only cases which are not well grouped are those corresponding to the lowest concentration rates (i.e. 0.01, or in some case 0.03 mg/m$^3$). The method does not resolve different phytoplankton assemblages at such a low concentration and groups these cases together in the same cluster (see the multicolor cluster at the bottom of the central panel in Figure 10).

The results confirm the potential of using all the information contained in the derivative of hyperspectral normalized $R_n(\lambda)$ in comparison with the use of multispectral measurements or band ratios of discrete spectral values, and are in concordance with the results of Torrecilla et al. (2008b). The HCA cluster analysis based on the hyperspectral input information, with variable optical conditions, was able to automatically bring together assemblages corresponding to the same phytoplankton groups. Furthermore, the performance of this derivative analysis tool is dependent on the examination of the spectral data and on the selection of suitable parameters (i.e. smoothing filter size and band separation) for each particular data set and for each specific purpose of the analysis. If smoothing and derivative analysis parameters are, for instance, selected too coarse, spectral features of interest will be lost and worse results will be achieved in the automatic identification of phytoplankton assemblages using clustering analysis of derivatives of $R_n(\lambda)$ spectra. This is the case shown in Figure 10 (bottom panel), where it can be seen that Prymnesiophyceae and diatoms are grouped together, and spectra corresponding to the group of Cyanophyceae (in blue) appear mixed with other phytoplankton groups.
Fig. 10. Cluster analyses based on: (top) raw hyperspectral normalized $R_{rs}(\lambda)$, (centre) second derivative of hyperspectral normalized $R_{rs}(\lambda)$ spectra and (bottom) second derivative of hyperspectral normalized $R_{rs}(\lambda)$ spectra using analyzing parameters that are too coarse (i.e. derivative band separation and smoothing filter size).

5.3 Impact of the sensor’s response on the proposed automatic identification

A sensor sensitivity analysis was also conducted to test the effect of the sensor’s response, with which the radiometric measurements would hypothetically be acquired in the proposed methodology (Figure 8) to identify different phytoplankton assemblages from derivatives of hyperspectral $R_{rs}(\lambda)$ spectra.

The response of a miniature hyperspectral CMOS-array spectrometer (Figure 3) designed for oceanographic monitoring purposes (Pons et al., 2007; Torrecilla et al., 2008a) was characterized in terms of sensitivity. The spectral sensitivity curve obtained was applied to both simulated hyperspectral water-leaving radiance ($L_w(\lambda)$) and downwelling irradiance ($E_d(\lambda)$), the ratio of which was used to estimate the new data set of $R_{rs}(\lambda)$ spectra. Each new $R_{rs}(\lambda)$ spectrum (Figure 11, red curve in the left panel) now has a spectral resolution of 3 nm approximately, in contrast to the 1 nm resolution of Hydrolight-modeled $R_{rs}(\lambda)$, since it has been adapted to the sensor’s spectral resolution. The magnitude of $R_{rs}(\lambda)$ spectra has not been essentially modified after the consideration of the sensor’s spectral sensitivity. This is because remote-sensing reflectance is an apparent optical property obtained as the ratio of two radiometric properties. This rationing serves as an effective way to remove the effects of the magnitudes of $E_d(\lambda)$, $L_w(\lambda)$, and other possible external light fluctuations. However, the small differences between $R_{rs}(\lambda)$ spectra with 1 and 3 nm spectral resolution become more evident on the results of derivative analysis (Figure 11, right panel) and will play an important role on the further clustering analysis.

Analogously to the previous subsection, derivative and clustering analysis was carried out with this new data set of hyperspectral-sensor $R_{rs}(\lambda)$. As can be seen in Figure 12, a proper identification and clustering of similar phytoplankton assemblages was only possible when derivative spectra were considered (right panel). Again, a cluster was also created by grouping all those cases corresponding to low concentrations (0.01 mg/m$^3$).
5.3 Impact of the sensor’s response on the proposed automatic identification

A sensor sensitivity analysis was also conducted to test the effect of the sensor’s response, with which the radiometric measurements would hypothetically be acquired in the proposed methodology (Figure 8) to identify different phytoplankton assemblages from derivatives of hyperspectral $R_{rs}(\lambda)$ spectra.

The response of a miniature hyperspectral CMOS-array spectrometer (Figure 3) designed for oceanographic monitoring purposes (Pons et al., 2007; Torrecilla et al., 2008a) was characterized in terms of sensitivity. The spectral sensitivity curve obtained was applied to both simulated hyperspectral water-leaving radiance ($L_w(\lambda)$) and downwelling irradiance ($E_d(\lambda)$), the ratio of which was used to estimate the new data set of $R_{rs}(\lambda)$ spectra. Each new $R_{rs}(\lambda)$ spectrum (Figure 11, red curve in the left panel) now has a spectral resolution of 3 nm approximately, in contrast to the 1 nm resolution of Hydrolight-modeled $R_{rs}(\lambda)$, since it has been adapted to the sensor’s spectral resolution. The magnitude of $R_{rs}(\lambda)$ spectra has not been essentially modified after the consideration of the sensor’s spectral sensitivity. This is because remote-sensing reflectance is an apparent optical property obtained as the ratio of two radiometric properties. This rationing serves as an effective way to remove the effects of the magnitudes of $E_d(\lambda)$, $L_w(\lambda)$, and other possible external light fluctuations. However, the small differences between $R_{rs}(\lambda)$ spectra with 1 and 3 nm spectral resolution become more evident on the results of derivative analysis (Figure 11, right panel) and will play an important role on the further clustering analysis.

Analogously to the previous subsection, derivative and clustering analysis was carried out with this new data set of hyperspectral-sensor $R_{rs}(\lambda)$. As can be seen in Figure 12, a proper identification and clustering of similar phytoplankton assemblages was only possible when derivative spectra were considered (right panel). Again, a cluster was also created by grouping all those cases corresponding to low concentrations (0.01 mg/m$^3$).
Fig. 12. Results of analog derivative and cluster analysis when the hyperspectral sensor’s response in terms of sensitivity is considered based on: (left) raw hyperspectral-sensor normalized R_{rs}(\lambda) and (right) second derivative of hyperspectral-sensor normalized R_{rs}(\lambda) spectra.

Though similar results were obtained, even including the effect of the sensor’s response in terms of sensitivity, it should be pointed out that the applied derivative spectroscopy needed to be adapted. Suitable values of smoothing filter size and derivative band separation were selected because of the different spectral resolution of the new R_{rs}(\lambda) data set. Furthermore, based on the results from this experiment, the selected low-cost and miniature hyperspectral CMOS-array spectrometer (Pons et al., 2007) has been confirmed as a potential tool for water component detection and monitoring.

6. Conclusion and Future Work

Derivative spectroscopy has been satisfactorily applied to numerical simulations of hyperspectral remote-sensing reflectance (R_{rs}(\lambda)) corresponding to different open-ocean environments. Its feasibility for identifying phytoplankton pigment assemblages in comparison with the use of raw R_{rs}(\lambda) spectra or traditional ratios of discrete spectral bands (Torrecilla et al., 2008b) has been confirmed using a validation approach based on hierarchical cluster analysis (HCA). Simulation-based experiments have also been performed, including a model of the response in terms of sensitivity of a hyperspectral sensor, with which radiometric measurements would hypothetically be acquired. The aim was to analyze the sensor’s effect on the proposed methodology of analysis. For each simulated hyperspectral data set, a proper adaptation of parameters involved in derivative analysis (i.e. smoothing filter size and derivative band separation) was necessary according to the characteristics of the signals and particularly their spectral resolution.

The experiments yield promising results when all the information contained in the second derivative of hyperspectral R_{rs}(\lambda) spectra is considered. This method can therefore provide a means for optical oceanographers to better characterize complex oceanic waters, detect harmful algal blooms or map phytoplankton functional types from hyperspectral oceanographic information. The recent advances in hyperspectral technology (e.g.
miniaturization and power supply reduction) have given rise to a great number of sensor configurations that are suitable for incorporating in a large number of in situ and remote sensing platforms of oceanographic observing systems (e.g. satellites, gliders), which will allow these challenges to be overcome once hyperspectral data become more available.

Future research will focus on experiments based on larger field hyperspectral data sets, since the potential of integrating hyperspectral $R_{\text{rs}}(\lambda)$ measurements and derivative spectroscopy has been emphasized in this work. However, the design of a more realistic approach when simulated-based experiments are employed would be useful for better validating the efficacy of the proposed method. An effort of detailed consideration of several the above-mentioned factors involved in the analysis process is suggested. For instance, the possibility of having some variability of the inherent optical properties (IOPs) used as an input in the radiative transfer modeling should be explored. Furthermore, the accuracy of the derivative-based method for identifying different phytoplankton compositions would be improved if distortion experiments, caused by the sensor (noise, spectral stray-light, thermal effects, etc.) were included in the simulation-based approach.

7. Acknowledgements

This study was supported by the projects HIDRA (PIE06-301102) and ANERIS (PIF08-015) funded by the Spanish Ministry of Science and Innovation.

8. References


Remote sensing is the acquisition of information of an object or phenomenon, by the use of either recording or real-time sensing device(s), that is not in physical or intimate contact with the object (such as by way of aircraft, spacecraft, satellite, buoy, or ship). In practice, remote sensing is the stand-off collection through the use of a variety of devices for gathering information on a given object or area. Human existence is dependent on our ability to understand, utilize, manage and maintain the environment we live in - Geoscience is the science that seeks to achieve these goals. This book is a collection of contributions from world-class scientists, engineers and educators engaged in the fields of geoscience and remote sensing.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
