CORE
Metadata, citation and similar papers at core.ac.uk
Provided by University of Queensland eSpace

Limnol. Oceanogr., 57(4), 2012, 1233–1244
© 2012, by the Association for the Sciences of Limnology and Oceanography, Inc.
doi:10.4319/lo.2012.57.4.1233

Physical environments of the Caribbean Sea

Iliana Chollett,[a,b,*] Peter J. Mumby,[b,a] Frank E. Müller-Karger,[c] and Chuanmin Hu[c]

[a] Marine Spatial Ecology Lab, College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom
[b] Marine Spatial Ecology Lab, School of Biological Sciences, University of Queensland, Brisbane, Australia
[c] Institute for Marine Remote Sensing, College of Marine Science, University of South Florida, St. Petersburg, Florida

Abstract

The Caribbean Sea encompasses a vast range of physical environmental conditions that have a profound influence on the organisms that live there. Here we utilize a range of satellite and in situ products to undertake a region-wide categorization of the physical environments of the Caribbean Sea (PECS). The classification approach is hierarchical and focuses on physical constraints that drive many aspects of coastal ecology, including species distributions, ecosystem function, and disturbance. The first level represents physicochemical properties including metrics of satellite sea surface temperature, water clarity, and in situ salinity. The second level considers mechanical disturbance and includes both chronic disturbance from wind-driven wave exposure and acute disturbance from hurricanes. The maps have a spatial resolution of 1 km^2. An unsupervised neural network classification produced 16 physicochemical provinces that can be categorized into six broad groups: (1) low water clarity and low salinity and average temperatures; (2) low water clarity but average salinity and temperature, broadly distributed in the basin; (3) low salinity but average water clarity and temperature; (4) upwelling; (5) high latitude; and (6) offshore waters of the inner Caribbean. Additional mechanical disturbance layers impose additional pattern that operates over different spatial scales. Because physical environments underpin so much of coastal ecosystem structure and function, we anticipate that the PECS classification, which will be freely distributed as geographic information system layers, will facilitate comparative analyses and inform the stratification of studies across environmental provinces in the Caribbean basin.

It has long been recognized that marine environments encompass vast environmental heterogeneity across a continuum of scales. Setting boundaries to the ocean is the first step towards its quantitative study (Longhurst 2007). In principle, a categorization of the physical marine environment might help explain patterns of the structure and function of marine systems and account for commonalities or contradictions in the results of experiments or monitoring studies carried out at different locations (Iken et al. 2010). Environmental classifications should also provide a logical means of stratifying field measurements so that outcomes can be scaled up appropriately.

Several attempts have been made to categorize the world's oceans into regions. The U.S. National Oceanic and Atmospheric Administration (NOAA) identified 64 large marine ecosystems: large areas with distinct bathymetry, hydrography, and productivity (Sherman and Hempel 2008). Longhurst (2007) classified the world ocean into 57 biogeochemical provinces after examining imagery of sea surface chlorophyll concentration and reviewing physical oceanographic literature for each ocean basin. Spalding et al. (2007) produced an expert-derived classification of the marine environment into 12 marine realms, 62 provinces, and 232 ecoregions, areas that are expected to hold a relatively homogeneous composition of species. These categorizations partition the Caribbean Sea into either two (Longhurst 2007; Sherman and Hempel 2008) or nine units (Spalding et al. 2007: fig. 1).

Neither of these classification schemes was intended to categorize the physical environment at an intra-Caribbean scale. This point is made clear by considering regional variability of sea surface temperature, an important environmental variable in determining pattern and function in marine systems (Tittensor et al. 2010). There is great thermal variability (Fig. 1) within even the most detailed classification available (Spalding et al. 2007), which did not take into consideration oceanographic data to carry out the categorization. For example, the southern Caribbean ecoregion (ecoregion 66 in Fig. 1) encloses areas influenced by upwelling (Müller-Karger et al. 1989) as well as offshore oligotrophic areas. Furthermore, similar physical environments have been arbitrarily separated into different ecoregions, as highlighted by the division of the Colombian upwelling areas (described by Andrade and Barton 2005) into ecoregions 66 and 67 (Fig. 1).

The physical environment of the Caribbean Sea is spatially heterogeneous, and such variations are likely to affect the function and distribution of marine organisms in the basin. Major sources of heterogeneity include river plumes (Müller-Karger et al. 1989; Restrepo et al. 2006), runoff (Imbach et al. 2010), upwelling (Müller-Karger et al. 2004; Andrade and Barton 2005), and bathymetric effects (Cerdeira-Estrada et al. 2005). These mechanisms principally alter the physicochemical environment experienced by marine organisms (i.e., temperature, light, salinity), which influence fundamental biological processes including metabolism and photosynthesis. None of the existent attempts to classify the ocean into regions (Longhurst 2007; Spalding et al. 2007; Sherman and Hempel 2008) capture this variability.

Here, we begin by creating a classification of the physicochemical environments of the Caribbean. This

* Corresponding author: i.c.chollett-ordaz@exeter.ac.uk

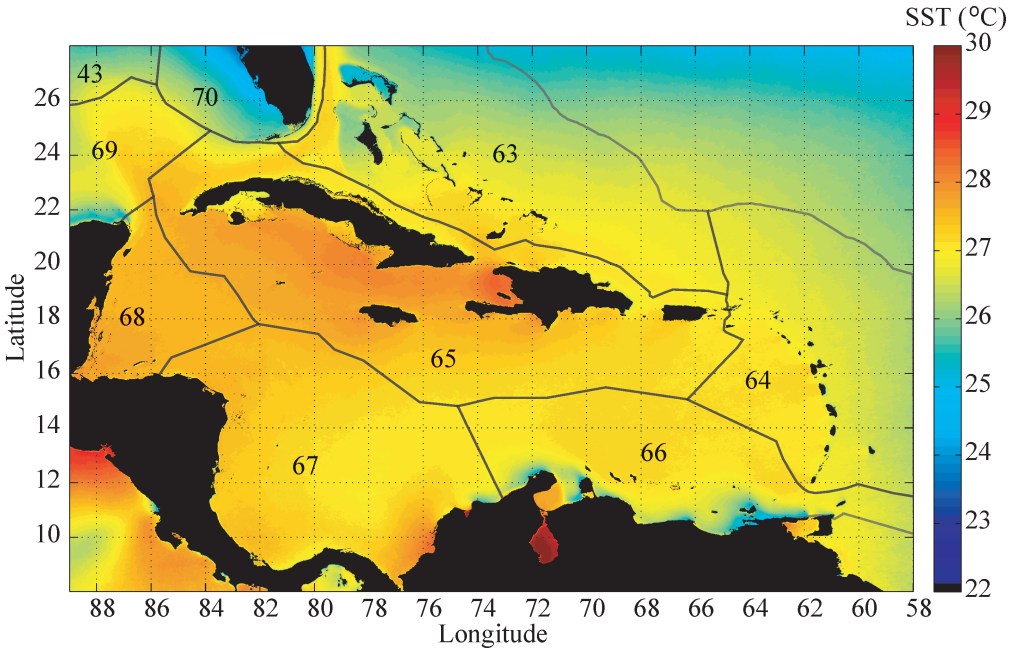

Fig. 1. Average sea surface temperature map (AVHRR 1993–2008) and marine ecoregions in the Caribbean Sea (8–28°N, 89–58°W) according to Spalding et al. (2007): (43) Northern Gulf of Mexico; (63) Bahamian; (64) Eastern Caribbean; (65) Greater Antilles; (66) Southern Caribbean; (67) South-western Caribbean; (68) Western Caribbean; (69) Southern Gulf of Mexico; (70) Floridian.

classification defines major physicochemical boundaries, but many organisms, particularly those living in shallow coastal habitats (< 20 m) are also strongly influenced by the mechanical disturbance regime. Two principal types of mechanical disturbance can be distinguished: chronic exposure to waves and acute, episodic physical disturbance from tropical cyclones (Connell 1978; Edwards et al. 2011). To accommodate physicochemical environments and different mechanical disturbances, we created a hierarchical classification scheme of the physical environments of the Caribbean Sea (PECS) that encompasses the fundamental physicochemical regime at the top level and two forms of physical disturbance at lower levels. The PECS classification was developed directly from observed data rather than being imposed upon data, and was implemented at a high spatial resolution of 1 km^2. We anticipate that the classification scheme will benefit the systematic quantitative study of biological oceanography in the Caribbean in addition to that of coastal habitats such as intertidal rocky shores, mangroves, subtidal seagrass beds, and coral reefs. Moreover, a detailed classification of physical environments will help inform conservation planning activities of the likely stratification of biodiversity in the area.

Methods

The data set—The region of interest covers the area between 8°N and 28°N and 89°W and 58°W. Waters of the Atlantic Ocean, 100 km away from the eastern edges of the land masses circumscribing the basin, were masked out and excluded from the analyses. The environment of the Caribbean Sea was defined in terms of sea surface temperature, water clarity, salinity, wind-driven wave exposure, and hurricane incidence. Surface physicochemical variables may differ from those measured at the subsurface (Leichter et al. 2006). The magnitude of these differences depends on local atmospheric and oceanographic conditions, particularly the level of stratification of the water column, which is spatially and temporally variable (Donlon et al. 2002). Although the input variables used here were measured at the ocean surface and therefore do not capture this vertical variability, superficial satellite measurements are the only ones that allow a full assessment of horizontal heterogeneity at a basin level. Additionally, surface temperature and ocean color are both strongly correlated to processes in the entire water column (Longhurst 2007; Oliver and Irwin 2008). Thus, the resulting classification has some value for the assessment of pelagic systems in addition to its main focus of benthic systems. A short description of the source data and the methodology followed to obtain each data layer is given below.

Sea surface temperature (SST) data were derived from infrared observations collected by the Advanced Very High Resolution Radiometer (AVHRR) sensors flown on the NOAA's Polar Orbiting Environmental Satellite Series (satellites NOAA 11–18). AVHRR data from 1993 to 2008 at 1-km^2 spatial resolution were gathered, navigated, processed, and archived by the Institute for Marine Remote Sensing (IMaRS) at the University of South Florida. The temporal coverage initiates when the High Resolution Picture Transmission antenna located at the University of South Florida first collected data. Nightly data were subjected to the cloud-filtering procedure described by Hu et al. (2009). In the shallow Florida Keys environment the root mean square difference of the 1-km AVHRR SSTs from those measured by several buoys was < 1°C (Hu et al.

2009), while in open-ocean environments the differences were smaller (Kearns et al. 2000). From this data set, we calculated the climatological average, the climatological minimum monthly mean (mMM), and the climatological maximum monthly mean (MMM). In order to calculate mMM and MMM, the average temperature for each month (i.e., the mean of all daily data from all years for that particular month) was calculated, and then the lowest and the highest values were selected to produce the mMM and MMM maps.

A proxy for *water clarity* was assessed using the diffuse attenuation coefficient at 490 nm (K490). K490 represents the rate which light in the blue to green region of the spectrum is attenuated with depth (i.e., higher values indicate lower water clarity). Time series data from 1997 to 2008 at 1-km^2 spatial resolution were derived from ocean color observations collected by the sea-viewing wide-field-of-view sensor (SeaWiFS) on board of GeoEye's SeaStar satellite. SeaWiFS data were collected, navigated, processed (using SeaDAS version 4 software, NASA's Ocean Biology Processing Group), and archived by IMaRS. Validation of the SeaWiFS default K490 data product over the global ocean (bottom depth > 30 m) through NASA's SeaWiFS bio-optical archive and storage system (SEABASS) online tool showed root mean square error of 40.1% ($n = 410$, $0.018 < \text{K490} < 0.58$, $R^2 = 0.77$) as compared with in situ measurements (http://seabass.gsfc.nasa.gov/seabasscgi/validation_search.cgi). A recent work using limited data collected in southwest Florida coastal waters and in the Caribbean showed root mean square uncertainty of 27.6% ($n = 7$, $0.03 < \text{K490} < 0.41$, $R^2 = 0.99$) (J. Zhao unpubl.). Saturated values (K490 = 6.3998 m^{-1}, the maximum attainable value for this variable) were removed from the data set. Spatial variability in the relative composition and nature of in-water constituents in coastal waters makes their quantification problematic in large regions that include both clear and optically turbid areas (Babin et al. 2003). Additionally, the Caribbean Sea contains optically shallow areas (e.g., the clear Bahamas banks) where light reflected by the seabed contributes to the reflectance spectra captured by the satellite sensor, resulting in misleadingly high K490 estimates (K490 values up to 0.4 m^{-1}, Cannizzaro and Carder 2006; J. Zhao unpubl.). Therefore, we calculated the relative frequency of K490 anomalies (percentage of times that K490 values were above 0.5 m^{-1}) as a proxy to identify areas of low water clarity throughout the Caribbean Sea. While the threshold of 0.5 m^{-1} accounts for the effects of optically shallow waters, it confines the approach to the detection of large water quality anomalies that occur only in coastal areas because it is unlikely to obtain K490 values above 0.5 m^{-1} in open-ocean waters. The use of relative frequencies allows the comparison of areas with different observation frequency because of heterogeneous cloud coverage in this large region.

Salinity data from the World Ocean Atlas 2009 were obtained from the NOAA National Oceanographic Data Center. The original data were collected from several sources, including bottle samples, ship-deployed conductivity–temperature–depth package, profiling float, moored and drifting buoys, gliders, and undulating oceanographic recorder profiles (Antonov et al. 2010). The earliest observations were recorded during the 17th century, and the last observations during 2008. The data set was analyzed in a consistent, objective manner on a 0.25° latitude–longitude grid at standard depth levels (Antonov et al. 2010). Here we used climatological monthly composites of salinity at the surface. From this source data we calculated average surface salinity, which was then rescaled to 1 km^2 using bicubic interpolation in order to match the spatial resolution of the other layers.

Wind-driven wave exposure, the degree of wave action on an open shore, is governed by the fetch (distance of open sea that the wind has crossed to generate waves) and the strength and direction of the winds. The approach used here, based in wave theory, excludes the influence of any other effects on the wave climate (i.e., tides, and swell arising from distant sources) that are not generated by the local wind. Although an approximation of wave patterns in shallow areas, simple methods based on the configuration of the coastline and wind patterns have repeatedly shown to be sufficient predictors of spatial variation in coastal communities (Chollett and Mumby 2012; Harborne et al. 2006). Here we measured fetch using the global, self-consistent, hierarchical, high-resolution shoreline database (GSHHS version 1.5, Wessel and Smith 1996), and wind speed and direction were acquired from the QuikSCAT (NASA) satellite scatterometer from 1999 to 2008. Ascending and descending passes were averaged in order to produce one daily wind estimate per day. Wind data, originally at ~ 25-km spatial resolution, were rescaled to 1 km^2 using bicubic interpolation prior to the analyses. Wave exposure was calculated using the method described by Ekebom et al. (2003), where the exposure of a location is a function of the shape of the basin, wind speed, and direction. However, the method of Ekebom et al. (2003) was developed in small archipelago environments (dozens of kilometers) with uniform wind conditions and fetch-limited exposure (Ekebom et al. 2003). For this study we are assessing a large region (thousands of kilometers) with high variability in wind distribution and many open, fetch-unlimited areas. For those reasons we made two modifications to the original method: By specifying the shift between equations for "fetch-limited" and "fully developed" seas, because for a given wind speed and a long fetch there is a fixed height to which a wave can grow (Chollett and Mumby 2012) and by including spatial variability in wind fields using gridded wind data. Additionally, we calculated daily wave exposure and then produced an average for the entire time period, instead of using the average wind speed in each of the main directions (Ekebom et al. 2003). This approach allows inclusion of strong, sporadic winds, which have a disproportionate influence on resulting wave patterns and would be missed otherwise. A detailed description of the overall method and equations can be found in Chollett and Mumby (2012).

Hurricane incidence was measured using the Atlantic Hurricane data set (1851–2008), which tracks the location and intensity of the eye of tropical cyclones every 6 h (Jarvinen et al. 1984). We confined the analyses to storms that reached hurricane intensity (i.e., no tropical storms

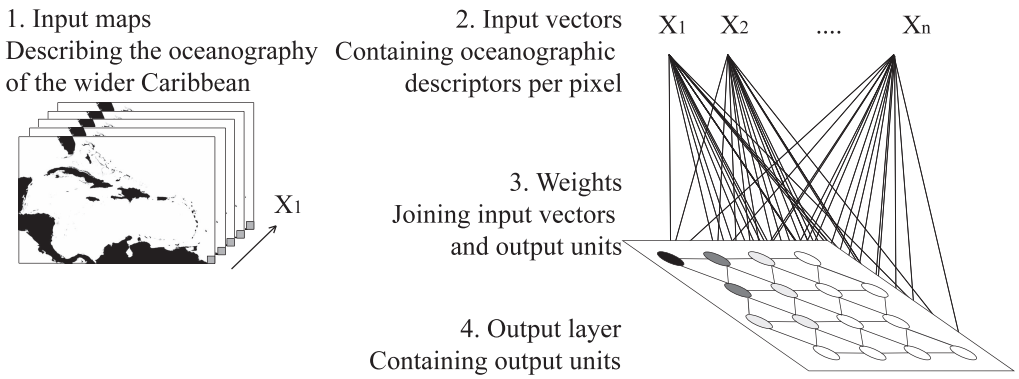

Fig. 2. Overview of the SOM algorithm: each pixel within the temperature (average, mMM, MMM), water clarity, and salinity input maps (1) is taken as an input vector for the SOM algorithm (2). *n* input vectors (where *n* is equal to the number of pixels in the image) are linked to the output layer through weighted links (3). The output layer (4) is, in this case, a matrix composed by 16 (4 × 4) units in a hexagonal grid, where lines indicate the connections among units, and the update neighborhood of the first (black) unit is defined by the gray gradient, with lighter colors highlighting farther units.

were included). Hurricane-force winds may extend several kilometers from the hurricane track. We calculated the frequency of hurricanes in any given location using the approach described by Edwards et al. (2011). Essentially, the area of influence of each hurricane is captured in buffers (up to 160 km wide) that take into account the intensity of the storm, its asymmetry (because of the Coriolis force), and the reduction in wind speed with distance from the hurricane track (Keim et al. 2007). Using this approach, we mapped the total frequency of hurricanes for each Saffir–Simpson intensity class for the entire record at 1-km^2 spatial resolution.

Classification—The Caribbean basin was classified into physicochemical regions using self-organizing maps (SOM; Kohonen 1995), an unsupervised clustering approach. The input variables for the classification were average SST, mMM SST, MMM SST, relative frequency of water clarity anomalies, and average surface salinity, which are expected to depict the most relevant physicochemical features of the area. All variables were first standardized to allow meaningful comparisons. For each variable, standardization was achieved by centering each value x_i (i.e., by subtracting the lowest score: $x_i - x_{min}$) and then dividing this by the range ($x_{max} - x_{min}$; Legendre and Legendre 1998: Eq. 1).

$$x_i' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

The classification method (SOM) is a type of neural network based on competitive learning that both reduces the dimensionality of the data and displays similarities among data. SOM was preferred over more traditional clustering approaches, such as hierarchical clustering or *k*-means, because of its ability to deal with large data sets and nonlinear problems. SOM does not make a priori assumptions about the distribution of the data, making it more appropriate for water clarity and salinity data, which have leptokurtic and skewed distributions. Other strengths of SOM are its ease of implementation, adaptation (the ability to change its structure based on external or internal information), parallelization (performing small operations in parallel), flexibility, and speed. A brief overview of the SOM algorithm is given below, but the reader is referred to Kohonen (1995) for a more technical discussion. Although the SOM has been used previously to extract patterns in satellite imagery (Richardson et al. 2003), most applications have employed it to extract temporal patterns, and, to our knowledge, only one study has used it for the extraction of spatial patterns and the identification of regions (Saraceno et al. 2006).

SOM is a nonlinear classification analysis in which multidimensional data are mapped onto a two-dimensional output space while preserving the topological relationships among input data (Kohonen 1995). The input data for the analysis are *N* vectors (one for each pixel in the study area, more than three million in total) with five dimensions, corresponding to the descriptors of the physicochemical environments of the Caribbean. The SOM (Fig. 2) consists of a set of units, nodes or neurons arranged in a two-dimensional grid. These units characterize the center of the clusters. The number of units (and clusters) and the type of arrangement (e.g., hexagonal or rectangular grid) are defined by the user and are dependent upon the level of detail desired in the analysis. A weight vector of the same dimension as the input data is associated with each unit. This vector is initialized with random values or eigenvectors of the data set. During the self-organizing process, input vectors are presented to the SOM and the distance of the weight vector of each unit to the input vector is calculated. The unit with the smallest distance is selected as the "winner." At this point, the weights of both the winner and its neighboring units are modified to more closely resemble the input vector. These changes depend on a learning rate, a constant which determines how fast the SOM learns and decreases with the number of iterations (units change more at the beginning of the iterative process) and a neighboring function that is spatially explicit (neighbor units farther away from the winner change less). The procedure is repeated until each input vector is presented to the network, and then the entire process is

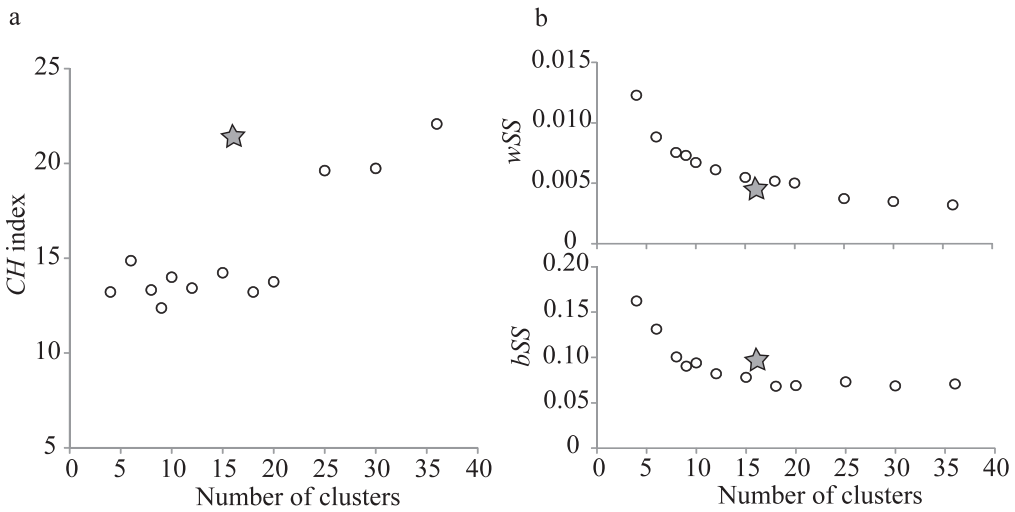

Fig. 3. Metrics used to identify the optimum number of clusters (a) Caliński and Harabasz index (*CH*); (b) within-cluster sum of squares (*wSS*) and between-cluster sum of squares (*bSS*). Gray stars correspond to 16 clusters, the partition chosen.

repeated many times, leading to a topologically ordered map. The inclusion of a neighborhood function implies that similar patterns are mapped onto neighboring regions on the map, while dissimilar patterns are mapped farther apart. Once the underlying patterns have been characterized with the output units, SOM results can be used to classify the input vectors into clusters, where each vector is represented by the most similar unit.

The SOM requires the user to define the desired number of clusters a priori. To identify the optimum number of clusters, we partitioned the data set using 4–36 clusters and compared the results using a validation criterion. While a few clusters produce groups that are well separated but internally very variable (highly dispersed), too many clusters give more compact but overlapping clusters. To choose the optimal number of clusters (k), we evaluated the

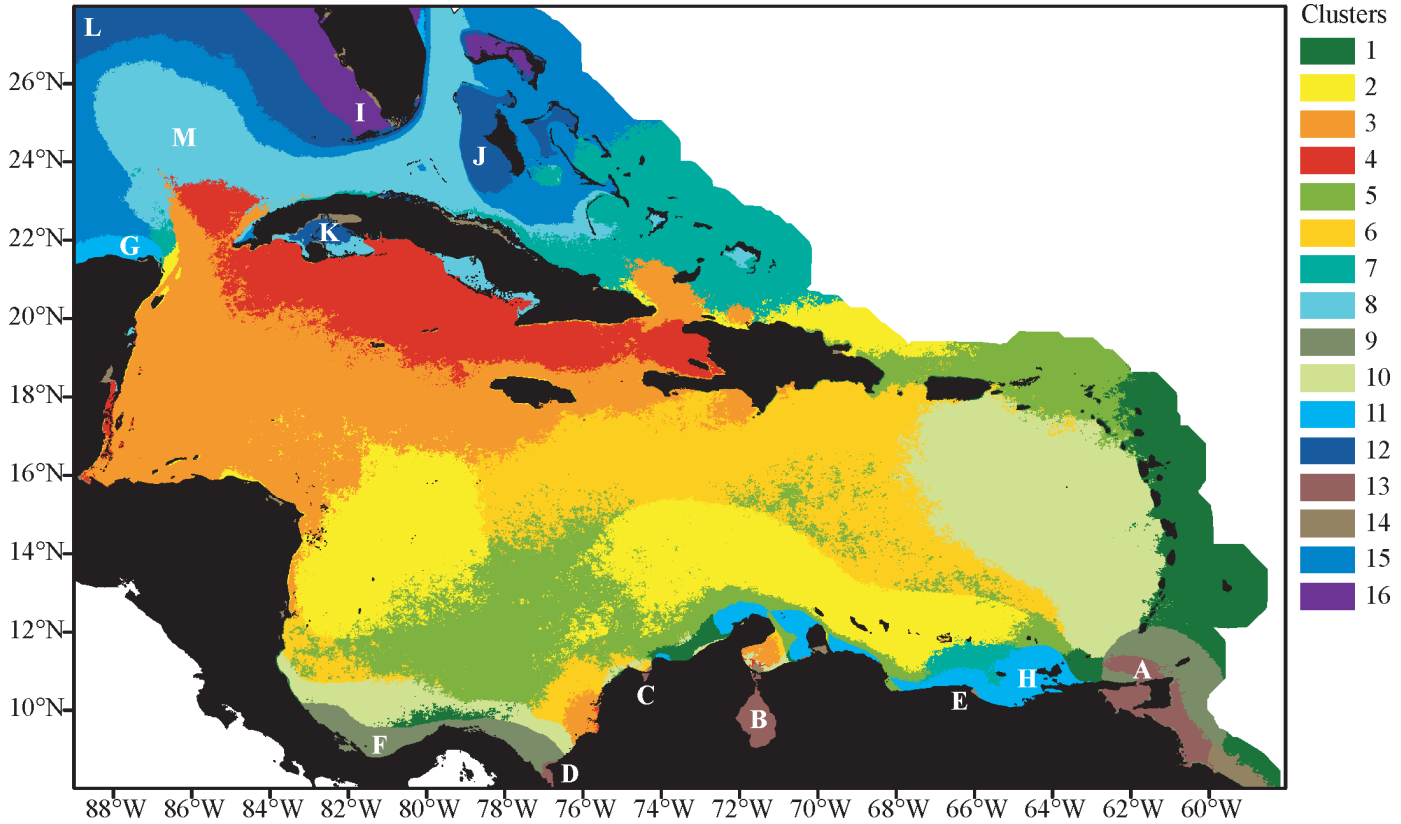

Fig. 4. Spatial arrangement of the 16 physicochemical provinces in the Caribbean Sea and selected examples (A–N), described in Table 2.

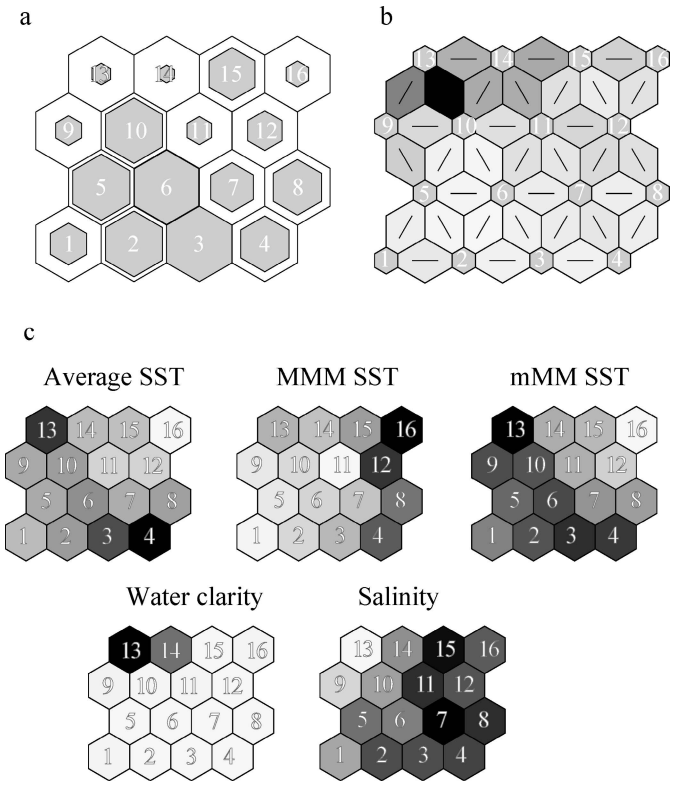

Fig. 5. Topological arrangement of the SOM showing the relative location of the clusters. (a) The proportion of pixels associated with each numbered cluster is indicated by the size of each hexagon, where larger hexagons indicate a larger number of pixels represented by that cluster. (b) Euclidean distances between the center of neighboring clusters, where the hexagons represent the clusters, the lines connect neighboring clusters, and the shades of gray in the regions containing the lines indicate distances between clusters. Darker shades represent larger distances and lighter colors smaller distances. (c) Average value for each cluster for each environmental variable. Darker shades represent larger values.

resulting classifications regarding their compactness or within-cluster variability, and isolation or between-cluster variability using the index described by Caliński and Harabasz (1974). The approach is analogous to the *F*-statistic in univariate analysis and has shown a superior performance when compared to other indices (Milligan and Cooper 1985). Within-cluster variability (*wSS*) was estimated by calculating the sum of squares of the Euclidean distances between each pixel (*n* pixels in total) and the centroid of its cluster, while between-cluster variability (*bSS*) was assessed as the sum of squares of the Euclidean distances between each cluster centroid and the centroid of the entire data set. The Caliński and Harabasz (*CH*) index is given by the formula below (Eq. 2). The number of clusters that maximizes *CH* suggests the best partition.

$$CH_k = \frac{\dfrac{bSS_k}{(k-1)}}{\dfrac{wSS_k}{(n-k)}} \tag{2}$$

The units of the SOM were arranged in a hexagonal grid, which allows a better visualization and more continuous transitions among the units. The learning rate was 0.9 during the initial phase and 0.02 during the refining phase. The initial neighborhood size was set to 3, and the refining-phase neighborhood size was set to 1. The initial phase consisted of 100 steps, while the refining phase consisted of 400 steps, for 500 iterations in total. To calculate distances from a particular unit to its neighbors we used the link distance, which is simply the number of links that must be taken to reach the unit under consideration. The analyses were performed using the Neural Network toolbox in Matlab version 7.10 (The MathWorks, Inc).

Geographic information system files with the PECS classification including both the physicochemical categorization and the physical disturbance regime are available from the corresponding author.

Table 1. Average and standard deviation of average SST, mMM, MMM, water clarity proxy (WCp), and salinity for each of the 16 clusters.

Cluster	Average SST (°C)	mMM SST (°C)	MMM SST (°C)	WCp (%)	Salinity (‰)
1	26.63±0.19	25.42±0.35	27.97±0.22	1.03±1.61	35.02±0.3
2	27.17±0.13	25.96±0.29	28.54±0.12	0.23±0.35	35.91±0.12
3	27.60±0.12	26.26±0.19	29.10±0.15	0.10±0.45	35.9±0.16
4	27.84±0.19	26.19±0.31	29.68±0.21	0.08±0.50	35.97±0.13
5	27.00±0.15	25.76±0.35	28.25±0.14	0.33±0.51	35.69±0.16
6	27.27±0.14	26.08±0.21	28.59±0.13	0.23±0.39	35.55±0.1
7	26.93±0.22	25.12±0.37	28.86±0.22	0.10±0.29	36.3±0.12
8	27.12±0.19	24.92±0.43	29.60±0.24	0.11±0.45	36.1±0.12
9	27.01±0.42	26.02±0.57	28.23±0.33	2.57±2.58	33.1±1.23
10	27.15±0.13	25.98±0.26	28.49±0.18	0.64±0.99	35.17±0.22
11	26.06±0.39	24.31±0.65	27.94±0.32	2.04±1.73	36.15±0.4
12	26.05±0.36	22.26±0.61	29.78±0.29	0.25±0.83	35.88±0.38
13	27.69±1.24	26.60±1.36	29.19±1.18	36.32±18.87	31.44±1.69
14	26.65±0.81	24.34±1.88	28.90±1.04	21.99±12.71	35.55±0.71
15	26.46±0.28	23.76±0.39	29.41±0.31	0.09±0.36	36.23±0.2
16	25.07±0.47	19.74±1.05	29.88±0.33	1.10±2.70	35.81±0.24

Table 2. Physicochemical provinces of the Caribbean Sea.

Main oceanographic features	Clusters	Example	Country	Description
Low water clarity, low salinity	13	A. Orinoco River plume	Venezuela	Under the freshwater influence of the Orinoco River, the fourth of the world's rivers in terms of discharge (Müller-Karger et al. 1989)
		B. Lake Maracaibo	Venezuela	The largest brackish lake in South America, influenced by the freshwater discharge of numerous rivers (Rodríguez 2000)
		C. Cienaga Grande de Santa Marta	Colombia	Large estuarine lagoon complex that forms part of the delta of the Magdalena River, the largest river discharging directly to the Caribbean Sea (Restrepo et al. 2006)
		D. Uraba Gulf	Colombia	The southernmost portion of the Caribbean Sea, with waters influenced by the freshwater discharges of the Atrato River and other small streams (Diaz et al. 2000)
Low water clarity	14	E. River Tuy	Venezuela	The River Tuy concentrates the wastewater effluents from the capital of Venezuela (Jaffe et al. 1995)
Low salinity	9	F. Panama–Costa Rica runoff region	Panama and Costa Rica	Under the influence of river discharge and runoff (Imbach et al. 2010) driven by strong rainfall in the area (Portig 1965)
Upwelling	11	G. Yucatan upwelling	Mexico	Topographically and wind-induced upwelling (Merino 1997; Melo-González et al. 2000)
		H. Southern Caribbean upwelling	Colombia and Venezuela	Wind-driven upwelling occurs along eastern Colombia and most of the Venezuelan coastline, although the Guajira (Andrade and Barton 2005) and the southeastern Venezuela (Müller-Karger et al. 2004) are the best known upwelling areas in the region
High-latitude areas	12,15,16	I. Florida banks	USA	Topographically induced fronts with colder waters in winter due to the action of sensible heat and evaporative losses in shallow areas
		J. Northern Bahamas banks	The Bahamas	Topographically induced fronts, with thermal contrasts enhanced at higher latitudes
		K. Western Cuba banks	Cuba	These topographically induced fronts have been described by Cerdeira-Estrada et al. (2005)
		L. Inner Gulf of Mexico	Mexico and USA	The limits of the waters of the inner Gulf of Mexico are clearly delineated by the loop current. Within the Gulf waters exhibit higher seasonal variation in temperature (Müller-Karger et al. 1991)
Inner Caribbean	1–8,10	N. The Loop Current	USA and Cuba	Joining the Yucatan and Florida currents in a clockwise flow (Hofmann and Worley 1986)

Results

SOM classification—The optimal partition of the input data set was found when using 16 clusters (Fig. 3). For this partition the CH index attained the second maximum value after 36 classes (Fig. 3a). This classification scheme represents the best tradeoff in minimizing the within-cluster variability (*wSS*) and maximizing the between-cluster variability (*bSS*, Fig. 3b) while providing a number of clusters that is small enough to be easily visualized and interpreted. Punctuated increases in CH index when using 16 clusters are most likely due to the bidimensional configuration of this network. Unlike other clustering methods, such as *k*-means, which looks for a partition just based on a given number of clusters, SOM arranges this fixed number of clusters in a topologically ordered partition. For example, the 16-cluster solution derives from a topologically ordered grid of 4×4 units (*see* Fig. 2),

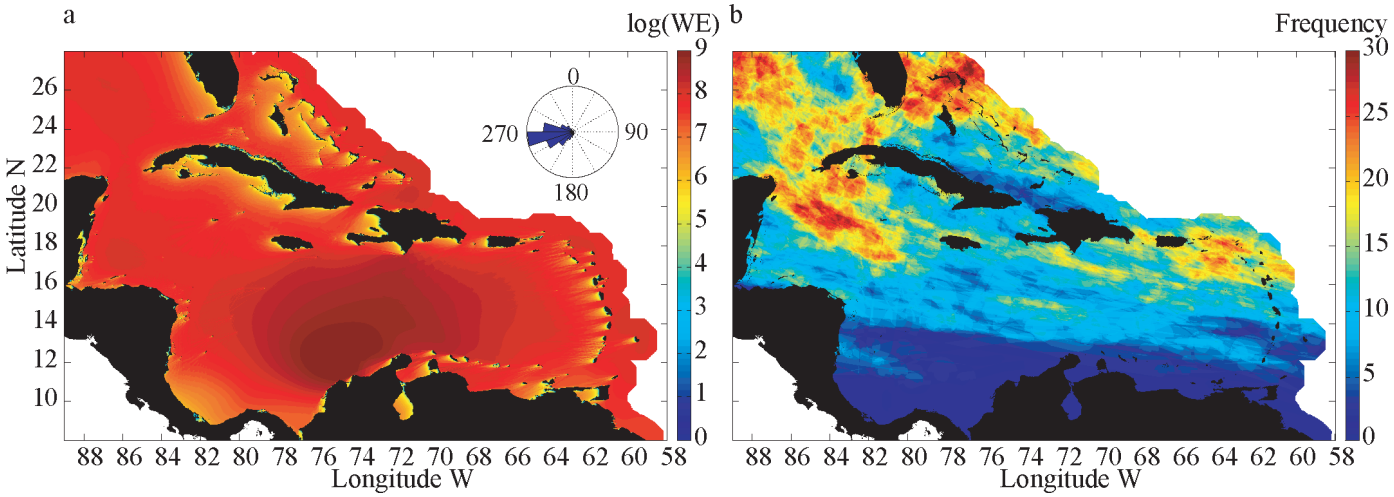

Fig. 6. Chronic and acute physical disturbance regime. (a) Chronic stress given by wave exposure (WE, in J m^{-3}). The wind rose in the top right corner shows the average wind conditions (1999–2008) for the entire basin. (b) Acute stress given by the frequency of occurrence of hurricanes Category 1–5 in the last 157 yr (1851–2008).

where all neighbor units are related. The CH index not only identifies the optimal number of classes but the optimal arrangement of the units, which is relevant for neighborhood-sensitive classification tasks.

The selected SOM classified the physicochemical environment of the Caribbean Sea into 16 clusters (Fig. 4). Although no explicit geographic constraints were included, the clustering procedure mostly produced homogenous clusters with well-defined boundaries. Even though maps obtained with larger numbers of clusters provide more complex patterns, the distribution of the main physicochemical features remained stable when different partitions were applied to the data set (not shown).

The number of pixels within each class was not evenly distributed and a few clusters (e.g., cluster 3, cluster 6) included most of the pixels, while others (e.g., cluster 13, cluster 14) enclosed a small subset of the region (Fig. 5a). The SOM (Fig. 5b) shows similar patterns adjacent of one another, dissimilar patterns at opposite ends of the SOM space, and a continuum of change across the array (Kohonen 1995). Clusters 13 and 14, on the top left corner, are very different from the rest, while clusters 2, 5, 6, and 10 are quite similar. These similar clusters represent offshore waters, with similar oceanographic characteristics, weak gradients, and fuzzy boundaries (Fig. 4). When looking at the relative importance of the different variables in each cluster (Table 1), it can be seen that cluster 13 is characterized by the lowest water clarity and the lowest salinity. Cluster 14 also shows low water clarity. Cluster 16 exhibits the lowest SST average and mMM and the highest MMM, while cluster 4 shows the highest average SST.

Physicochemical provinces of the Caribbean Sea—The physicochemical provinces of the Caribbean Sea can be broadly distributed into six groups, which are detailed below and in Table 2. (1) Areas characterized by low water clarity and low salinity, but average temperatures were identified by the cluster 13. Examples are the Orinoco River plume (region A in Fig. 4), an area under the direct influence of the Orinoco River; the Lake Maracaibo (region B); Cienaga Grande de Santa Marta (region C); and the Uraba Gulf (region D).

(2) Areas characterized by low water clarity, but not exceptionally low values of salinity (cluster 14) and average temperatures were broadly distributed in the Caribbean Sea. These conditions exist, for example, in the Golfete de Coro (Venezuela), and in areas influenced by the Tuy (Venezuela, Region E in Fig. 4) and Magdalena (Colombia) river plumes. In Central America, turbid areas are located along the coast of Nicaragua and eastern Honduras, the Gulf of Honduras, Chetumal and Espiritu Santo Bays, and in the Conil lagoon in the north of Quintana Roo, Mexico. In North America, turbid areas are located in Tampa Bay, Charlotte Harbor, and Florida Bay. Finally, the Gulf of Batabano, eastern La Juventud Island, and the Camagüey Archipelago in Cuba and Samana Bay in Dominican Republic have high turbid areas in the Greater Antilles.

(3) Areas characterized by low values of salinity (cluster 9) but relatively high water clarity when compared to Provinces 1 and 2, were located along the coast of Panama and Costa Rica (region F in Fig. 4) and on the edge of the Orinoco River plume.

(4) Areas with the lowest seasonal temperature maximum, also characterized by generally cold average and minimum SSTs (cluster 11) were located in the upwelling areas of Yucatan (region G in Fig. 4) and the southern Caribbean (region H).

(5) Towards the north, areas with low average and minimum temperature but high seasonal maxima were characterized by clusters 12, 15, and 16. Shallower sections such as the Florida banks (region I in Fig. 4), the northern Bahamas banks (region J), and the western Cuban banks (region K) were characterized by larger seasonal ranges when compared to surrounding areas, experiencing colder waters in winter and warmer waters in summer. Cold waters of the inner Gulf of Mexico (region L) also share a similar temperature signature with broad seasonal ranges.

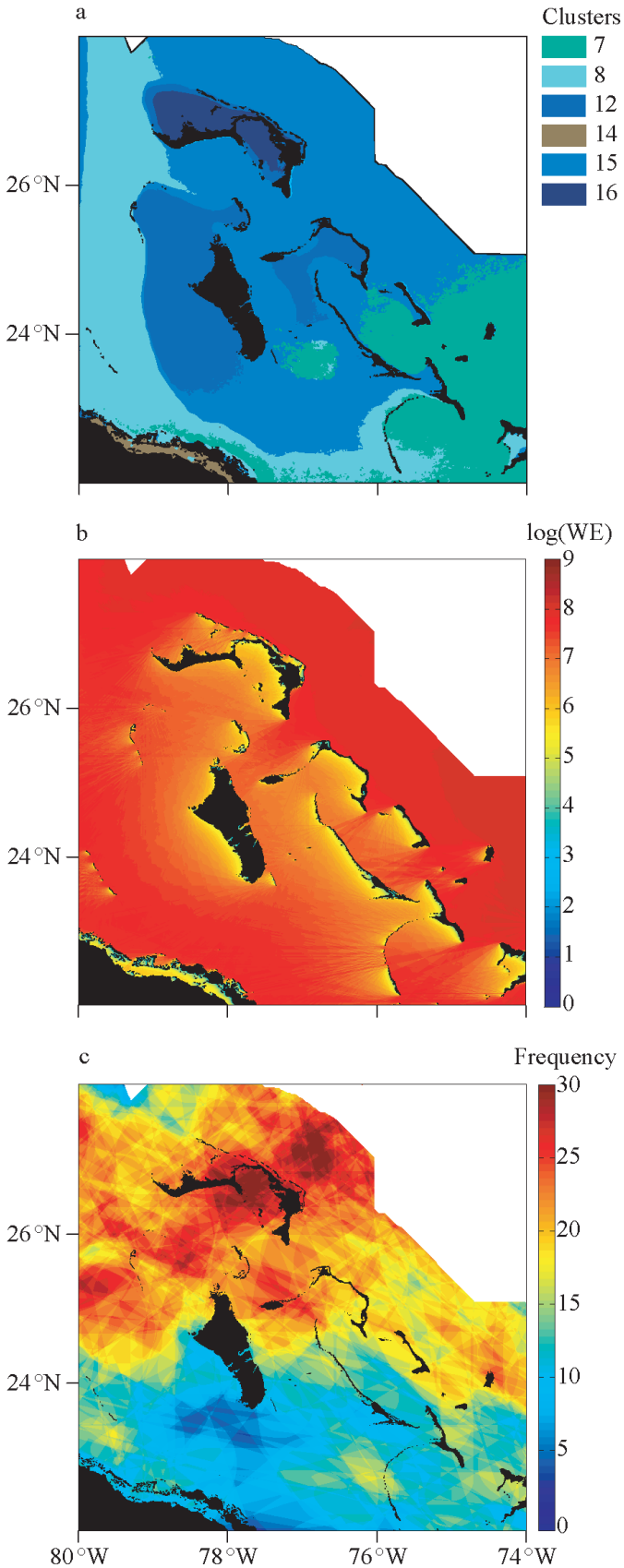

Fig. 7. Case study of northern Bahamas. (a) Physicochemical provinces, (b) chronic, and (c) acute physical disturbance regime.

(6) The interior of the Caribbean Sea was characterized by several (1–8, 10) highly correlated clusters (Fig. 5b). These clusters are characterized by a mixture of relatively warm waters with high salinity and high water clarity. Most differences within this area are fuzzy, indicating smooth transitions between clusters. The only exceptions are the waters of the loop current that identify a well-defined circulation path (region M).

Physical disturbance in the Caribbean Sea—To complement the oceanographic regime, the Caribbean Sea was characterized according to its chronic and acute physical disturbance regime (Fig. 6). Chronic physical disturbance, represented by wave exposure patterns, changes predictably in the basin according to the prevailing direction of the wind and fetch (Fig. 6a). The dominant effect of the northeasterly trade winds is clearly visible. In general, windward areas have higher wave exposure than leeward areas, unless they are sheltered by a land mass (e.g., westward cays of The Bahamas). From 1851 to 2008, 2199 hurricanes have been observed in the Caribbean basin. There is a clear spatial heterogeneity in the distribution of storms across the basin, with higher occurrence in the north and two centers of high activity in the passage between Yucatan and Cuba and east Puerto Rico (Fig. 6b).

The chronic and acute disturbance regime can be used to refine the physicochemical provinces if relevant for the system under study. The northern Bahamas, for example, encompasses three physicochemical provinces characterized by bathymetry-driven temperature effects (Fig. 7a), but this simple categorization can be enriched for shallow marine ecosystems by incorporating information on the chronic and acute disturbance regime. While chronic stress highlights spatial variability at the scale of a few kilometers (Fig. 7b), the acute disturbance regime in the area shows a marked latitudinal gradient (Fig. 7c).

Discussion

Information on the physicochemical characteristics of the water masses and the physical disturbance regime was used to produce the hierarchical PECS classification. While the physicochemical clusters are arranged in broad-scale spatial patterns, the physical disturbance imposes additional pattern that operates over fine (wave exposure) and large (hurricane incidence) spatial scales. The PECS classification is useful for research and conservation planning at regional scales by providing comprehensive coverage, a data-driven, objective classification approach, and high spatial detail consistent with the scale of many research and conservation requirements in the area. Although the data set is specific to the Caribbean and captures information relevant to the functioning of coastal ecosystems in the basin, the methodology applied is fully transferable to other geographic locations, other spatial scales, and can be used with other environmental variables.

The Caribbean Sea was divided into physicochemical regions characterized by similar sea surface water clarity, salinity, and temperature patterns. These regions delineate

Table 3. Issues that would benefit by referencing PECS.

Approach	Issue	Action
A priori	Selection of locations to facilitate comparisons	Choose locations from physicochemical provinces that are similar or contrasting according to the aims of the study
A posteriori	Selection of locations to place results into context	Compare results with locations within the same environmental province
A posteriori	Modeling patterns of marine ecosystem attributes at regional scales	Use environmental province as an explanatory variable
A posteriori	Priority setting and planning	Use environmental province as ecological criteria within selection algorithms

marine areas that are expected to hold similar ecosystem function and pattern. The regions, however, should not be interpreted as units with no interaction. The Caribbean Sea is highly connected, both in terms of physical oceanographic transport and the dispersive stages of many organisms (Cowen et al. 2006), which can lead to extensive gene flow across the region (Foster et al. 2012). Patterns of connectivity were not considered in this classification but would make for a useful extension, particularly for the study of marine biogeography within the region (Miloslavich et al. 2010). The inclusion of information on the connectivity among environmental provinces would make some of the boundaries defined in Fig. 4 fuzzier or, conversely, clearer, if oceanographic features contribute to the establishment of distinct ecological barriers (Cowen et al. 2006).

While some regions of the Caribbean have well-defined boundaries, indicating the presence of persistent features such as river and upwelling fronts, some regions in the inner Caribbean have fuzzy edges indicating transition areas. The resulting classification is consistent with physical oceanographic literature for the basin (Table 2). However, precise boundaries should be interpreted cautiously because they indicate *average* locations of rapid change in oceanographic conditions. Oceanographic features such as river plumes or upwelling fronts exhibit strong seasonal and interannual variability (Müller-Karger et al. 1989, 2004), and therefore have dynamic boundaries that are not captured by this static approach (Oliver and Irwin 2008). Additionally, some of these boundaries might be shifting under a changing climate. For example, changes in the strength of the upwelling in the southern Caribbean during the last few years have produced major changes in plankton communities in the area (F. E. Müller-Karger unpubl.). However, the extent to which recent changes in climate are actually shifting the ecotones and producing permanent changes in marine ecosystems is poorly understood.

Our objective was to produce a classification that can inform research and management goals that require spatial data on the structure and function of marine systems on timescales of at least 1 yr (e.g., structure of reef communities assessed in different parts of the Caribbean). If the desire is to compare patterns or processes that develop over short (e.g., a seasonal) timescales, such as the summertime community structure of plankton, then an alternative approach might be necessary. In this case, the same methodology could be used but the input data would be limited to include only the temporal scale of interest.

An important difference between PECS and earlier classification schemes is that PECS allows regions with similar oceanography to be defined. This enables regional comparisons of patterns or processes in often distant, but comparable systems. The categorization proposed here does not replace global classification systems (Longhurst 2007; Spalding et al. 2007; Sherman and Hempel 2008) but provides a more detailed regional product that fulfills regional research and conservation needs. Previous categorizations aim to provide enough detail to support linkage to applied research and conservation (Spalding et al. 2007); however, several research (Cruz-Motta et al. 2010; Iken et al. 2010; Miloslavich et al. 2010) and management (Mills et al. 2010) assessments have observed a mismatch in spatial scales between these zones and practical applications. Indeed, higher-resolution environmental data should better explain patterns of biodiversity and organismal distribution (Cruz-Motta et al. 2010; Iken et al. 2010), and smaller regions have been highlighted as more suitable for conservation activities because they have relatively homogeneous natural attributes, human activities, and aspects of governance that facilitate management (Mills et al. 2010).

The use of environmental classes is a common strategy in conservation planning (Ferrier 2002, Ferrier et al. 2009). The use of PECS depends on the level of environmental variability at the scale of interest. For example, regional planning activities, such as the identification of biodiversity hot spots (Harborne et al. 2006) or large-scale reserve networks, might identify hot spots of PECS diversity or seek to obtain representation of all PECS classes as part of a large-scale biodiversity plan (Mills et al. 2010). At smaller spatial scales, where only a limited number of PEC categories exist, practitioners might make use of the underlying data within physicochemical categories to stratify planning activities within individual PEC categories. PECS could also be used as a baseline against which to contrast spatial patterns of extreme weather. This would allow examining the relationship between climatological and stressful conditions, which varies spatially and is known to affect the responses of marine organisms to stress events (Mumby et al. 2011). Other potential applications of PECS include the production of hypotheses about geographic patterns of biological distinctiveness, including differences in composition, structure, function, and acclimation.

PECS facilitates a number of research and conservation applications, regardless if the researcher is in the planning stage or has already collected the data (Table 3). In general, the hierarchical approach provides an objective framework in which to plan, analyze, and interpret research and/or conservation efforts in the basin.

Acknowledgments

Funding for this effort was provided by the European Union Future of Reefs in a Changing Environment (FORCE) project, the National Aeronautics and Space Administration (NASA) Ocean Biology and Biogeochemistry program, Gulf of Mexico program, and Ocean Biodiversity program. We thank the Fondo Nacional de Ciencia, Tecnología e Innovación (FONACIT) of Venezuela, the Carbon Retention In A Colored Ocean (CARIACO) Ocean Time-Series Program (National Science Foundation grant 0963028) and NASA grant NNX09AV24G to FMK. We are grateful to Megan Saunders (University of Queensland) and two anonymous reviewers for valuable comments on the manuscript.

References

Andrade, C. A., and E. D. Barton. 2005. The Guajira upwelling system. Cont. Shelf Res. **25:** 1003–1022, doi:10.1016/j.csr.2004.12.012

Antonov, J. I., and others. 2010. World ocean atlas 2009, v. 2: Salinity, p. 184. *In* S. Levitus [ed.], NOAA atlas NESDIS 69. U.S. Government Printing Office.

Babin, M., D. Stramski, G. M. Ferrari, H. Claustre, A. Bricaud, G. Obolensky, and N. Hoepffner. 2003. Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. J. Geophys. Res. **108:** 3211, doi:10.1029/2001JC000882

Caliński, T., and J. Harabasz. 1974. A dendrite method for cluster analysis. Commun. Stat. **3:** 1–27, doi:10.1080/03610927408827101

Cannizzaro, J. P., and K. L. Carder. 2006. Estimating chlorophyll *a* concentrations from remote-sensing reflectance in optically shallow waters. Remote Sens. Environ. **101:** 13–24.

Cerdeira-Estrada, S., F. Müller-Karger, and A. Gallegos-Garcia. 2005. Variability of the sea surface temperature around Cuba. Gulf Mex. Sci. **2:** 161–171.

Chollett, I., and P. J. Mumby. 2012. Predicting the distribution of *Montastraea* reefs using wave exposure. Coral Reefs **31:** 493–503, doi:10.1007/s00338-011-0867-7

Connell, J. H. 1978. Diversity in tropical rain forests and coral reefs. Science **199:** 1302–1310, doi:10.1126/science.199.4335.1302

Cowen, P. K., C. B. Paris, and A. Srinivasan. 2006. Scaling of connectivity in marine populations. Science **311:** 522–527, doi:10.1126/science.1122039

Cruz-Motta, J. J., and others. 2010. Patterns of spatial variation of assemblages associated with intertidal rocky shores: A global perspective. PLoS ONE **5:** e14354, doi:10.1371/journal.pone.0014354

Diaz, J. M., G. Diaz-Pulido, and J. A. Sanchez. 2000. Distribution and structure of the southernmost Caribbean coral reefs: Golfo de Uraba, Colombia. Sci. Mar. **64:** 327–336, doi:10.3989/scimar.2000.64n3327

Donlon, C. J., P. J. Minnett, C. Gentemann, T. J. Nightingale, I. J. Barton, B. Ward, and M. J. Murray. 2002. Toward improved validation of satellite sea surface skin temperature measurements for climate research. J. Clim. **15:** 353–369, doi:10.1175/1520-0442(2002)015<0353:TIVOSS>2.0.CO;2

Edwards, H. J., and others. 2011. How much time can herbivore protection buy for coral reefs under realistic regimes of hurricanes and coral bleaching? Global Change Biol. **17:** 2033–2048, doi:10.1111/j.1365-2486.2010.02366.x

Ekebom, J., P. Laihonen, and T. Suominen. 2003. A GIS-based step-wise procedure for assessing physical exposure in fragmented archipelagos. Estuar. Coast. Shelf Sci. **57:** 887–898, doi:10.1016/S0272-7714(02)00419-5

Ferrier, S. 2002. Mapping spatial pattern in biodiversity for regional conservation planning: Where to from here? Syst. Biol. **51:** 331–363, doi:10.1080/10635150252899806

———, D. P. Faith, A. Arponen, and M. Drielsma. 2009. Community-level approaches to spatial conservation prioritization, p. 94–121. *In* A. Moilanen, K. A. Wilson, and H. P. Possingham [eds.], Spatial conservation prioritization: Quantitative methods and computational tools. Oxford Univ. Press.

Foster, N. L., and others. 2012. Connectivity of Caribbean coral populations: Complementary insights from empirical and modelled gene flow. Mol. Ecol. **21:** 1143–1157, doi:10.1111/j.1365-294X.2012.05455.x

Harborne, A. R., P. J. Mumby, K. Zychaluk, J. D. Hedley, and P. G. Blackwell. 2006. Modeling the beta diversity of coral reefs. Ecology **87:** 2871–2881, doi:10.1890/0012-9658(2006)87[2871:MTBDOC]2.0.CO;2

Hofmann, E. E., and S. J. Worley. 1986. An investigation of the circulation of the Gulf of Mexico. J. Geophys. Res. **91:** 14221–14236, doi:10.1029/JC091iC12p14221

Hu, C. M., and others. 2009. Building an automated integrated observing system to detect sea surface temperature anomaly events in the Florida Keys. IEEE Trans. Geosci. Remote **47:** 1607–1620, doi:10.1109/TGRS.2008.2007425

Iken, K., and others. 2010. Large-scale spatial distribution patterns of echinoderms in nearshore rocky habitats. PLoS ONE **5:** e13845, doi:10.1371/journal.pone.0013845

Imbach, P., L. Molina, B. Locatelli, O. Roupsard, P. Ciais, L. Corrales, and G. Mahe. 2010. Climatology-based regional modelling of potential vegetation and average annual long-term runoff for Mesoamerica. Hydrol. Earth Syst. Sci. **14:** 1801–1817, doi:10.5194/hess-14-1801-2010

Jaffe, R., I. Leal, J. Alvarado, P. Gardinali, and J. Sericano. 1995. Pollution effects of the Tuy River on the central Venezuelan coast: Anthropogenic organic compounds and heavy metals in *Tivela mactroidea*. Mar. Pollut. Bull. **30:** 820–825, doi:10.1016/0025-326X(95)00087-4

Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis. 1984. A tropical cyclone data tape for the north Atlantic basin, 1886–1983: Contents, limitations and uses [Internet]. NOAA Technical Memorandum NWS NHC 22. Miami (FL): National Hurricane Center [accessed 2011 January 1]. p. 21. Available from http://www.nhc.noaa.gov/pdf/NWS-NHC-1988-22.pdf

Kearns, E. J., J. A. Hanafin, R. H. Evans, P. J. Minnett, and O. B. Brown. 2000. An independent assessment of Pathfinder AVHRR sea surface temperature accuracy using the Marine-Atmosphere Emitted Radiance Interferometer (M-AERI). Bull. Am. Meteor. Soc. **81:** 1525–1536, doi:10.1175/1520-0477(2000)081<1525:AIAOPA>2.3.CO;2

Keim, B. D., R. A. Muller, and G. W. Stone. 2007. Spatiotemporal patterns and return periods of tropical storm and hurricane strikes from Texas to Maine. J. Clim. **20:** 3498–3509, doi:10.1175/JCLI4187.1

Kohonen, T. 1995. Self-organizing maps. Springer.

Legendre, P., and L. Legendre. 1998. Numerical ecology. Elsevier.

Leichter, J. J., B. Helmuth, and A. M. Fischer. 2006. Variation beneath the surface: Quantifying complex thermal environments on coral reefs in the Caribbean, Bahamas and Florida. J. Mar. Res. **64:** 563–588, doi:10.1357/002224006778715711

Longhurst, A. 2007. Ecological geography of the sea, 2nd ed. Academic Press.

Melo-González, N. M., F. E. Müller-Karger, S. Cerdeira-Estrada, R. Pérez De Los Reyes, I. Victoria Del Río, P. Cárdenas-Pérez, and I. Mitrani-Arenal. 2000. Near-surface phytoplankton distribution in the western Intra-Americas Sea: The influence of El Niño and weather events. J. Geophys. Res. **105:** 14029–14043, doi:10.1029/2000JC900017

Merino, M. 1997. Upwelling on the Yucatan shelf: Hydrographic evidence. J. Mar. Syst. **13:** 101–121, doi:10.1016/S0924-7963(96)00123-6

Milligan, G., and M. Cooper. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika **50:** 159–179, doi:10.1007/BF02294245

Mills, M., R. L. Pressey, R. Weeks, S. Foale, and N. C. Ban. 2010. A mismatch of scales: Challenges in planning for implementation of marine protected areas in the Coral Triangle. Conserv. Lett. **3:** 291–303, doi:10.1111/j.1755-263X.2010.00134.x

Miloslavich, P., and others. 2010. Marine biodiversity in the Caribbean: Regional estimates and distribution patterns. PLoS ONE **5:** e11916, doi:10.1371/journal.pone.0011916

Müller-Karger, F. E., C. R. Mcclain, T. R. Fisher, W. E. Esaias, and R. Varela. 1989. Pigment distribution in the Caribbean Sea: Observations from space. Prog. Oceanogr. **23:** 23–64, doi:10.1016/0079-6611(89)90024-4

———, R. Varela, R. Thunell, Y. Astor, H. Zhang, R. Luerssen, and C. Hu. 2004. Processes of coastal upwelling and carbon flux in the Cariaco Basin. Deep-Sea Res. II **51:** 927–943, doi:10.1016/j.dsr2.2003.10.010

———, J. J. Walsh, R. H. Evans, and M. B. Meyers. 1991. On the seasonal phytoplankton concentration and sea surface temperature cycles of the Gulf of Mexico as determined by satellites. J. Geophys. Res. **96:** 12645–12665, doi:10.1029/91JC00787

Mumby, P. J., and others. 2011. Reserve design for uncertain responses of coral reefs to climate change. Ecol. Lett. **14:** 132–140, doi:10.1111/j.1461-0248.2010.01562.x

Oliver, M. J., and A. J. Irwin. 2008. Objective global ocean biogeographic provinces. Geophys. Res. Lett. **35:** L15601, doi:10.1029/2008GL034238

Portig, W. H. 1965. Central American rainfall. Geogr. Rev. **55:** 68–90, doi:10.2307/212856

Restrepo, J. D., P. Zapata, J. A. Díaz, J. Garzón-Ferreira, and C. B. García. 2006. Fluvial fluxes into the Caribbean Sea and their impact on coastal ecosystems: The Magdalena River, Colombia. Global Planet. Change **50:** 33–49.

Richardson, A. J., C. Risien, and F. A. Shillington. 2003. Using self-organizing maps to identify patterns in satellite imagery. Prog. Oceanogr. **59:** 223–239, doi:10.1016/j.pocean.2003.07.006

Rodríguez, G. 2000. The Maracaibo system, Venezuela, p. 47-60. *In* U. Seeliger and B. Kjerfve [eds.], Coastal marine ecosystems of Latin America. Springer.

Saraceno, M., C. Provost, and M. Lebbah. 2006. Biophysical regions identification using an artificial neuronal network: A case study in the South Western Atlantic. Adv. Space Res. **37:** 793–805, doi:10.1016/j.asr.2005.11.005

Sherman, K., and G. Hempel. 2008. Perspectives on regional seas and the large marine ecosystem approach, p. 3–23. *In* K. Sherman and G. Hempel [eds.], 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 182. United Nations Environment Programme.

Spalding, M. D., and others. 2007. Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. BioScience **57:** 573–583, doi:10.1641/B570707

Tittensor, D. P., C. Mora, W. Jetz, H. K. Lotze, D. Ricard, E. V. Berghe, and B. Worm. 2010. Global patterns and predictors of marine biodiversity across taxa. Nature **466:** 1098–1101, doi:10.1038/nature09329

Wessel, P., and W. H. F. Smith. 1996. A global, self-consistent, hierarchical, high-resolution shoreline database. J. Geophys. Res. **101:** 8741–8743, doi:10.1029/96JB00104

Associate editor: James J. Leichter

Received: 20 December 2011
Amended: 08 May 2012
Accepted: 10 May 2012