Data Acquisition and Meta-analysis of Habitat Information from Gulf of Mexico Trophic Studies using CMECS

NOAA Contract # EA133C-05-CQ-1051

Fugro Earthdata Project # E09-0039-00

Final Report

Submitted by

Judy Soule, NatureServe

June 17, 2010

Principal Investigator

Dr. May Yuan, University of Oklahoma

Student Assistants

Lesley Williams

Sara Gonzalez-Perez

Derek Morris (funded by Oklahoma NASA Space Grant)

Consultant

Dr. James Simons, Texas Parks and Wildlife

Project starting date

September 1, 2009

Project ending date

May 14, 2010

Table of Contents

Project Introduction

Species habitats are fundamental to the understanding of ecosystem functions and consequently effective management of ecological resources. The Gulf of Mexico is notorious among fisheries biologists and resource managers for the poor state of data accessibility and incomplete biological understanding. The project presents an initial step towards the development of an ecological information infrastructure emphasizing species habitats in the Gulf of Mexico with the following two complementary aims:

> *Aim #1:* Test and revise the Coastal and Marine Ecological Classification Standard with fish habitat information in the Gulf of Mexico
>
> *Aim #2:* Build a sample spatial database of fish habitats in the Gulf of Mexico based on 400 published and unpublished studies within the estuarine and marine environments in the Gulf of Mexico, one of the 64 large marine ecosystems of the world.

Both aims contribute to larger scopes of research endeavors. The Coastal and Marine Ecological Classification Standard (CMECS) is intended to be the classification standard relevant to all U.S. coastal and marine environments and applicable to local, regional, and continental scales. CMECS provides a uniform protocol for identifying and naming ecological units with common coding schemes. In this study, Aim #1 was addressed through extraction of ecological data from research publications to examine classification standards and coding schemes as specified in CMECS III released in February 2009. Application of CMECS to ecological data from the literature illuminated practical issues that highlight needs for further CMECS improvements.

The GIS database of fish habitats developed in fulfillment of Aim #2 contributes to filling the ecological knowledge gap in the Gulf of Mexico. Ultimately, this study will become one piece of a cyber infrastructure of eco-informatics for species habitats and trophic dynamics in the Gulf of Mexico that will facilitate research of the spatiotemporal patterns in the habitat and tropic structure of macro- and mega-fauna in the large marine ecosystem. Combined with CMECS, the methodology we developed in this project is expected to be readily applicable to other marine ecosystems.

Methods and Materials

Ecological data used to test CMECS III were extracted from graduate theses, peer reviewed articles from primary literature, government reports, abstracts, and proceedings of meetings/conferences. The project consultant, Dr. James Simons, a marine ecologist at Texas Parks and Wildlife Department, provided the OU research team with copies of references on the habitats of fish in the Gulf of Mexico. These publications are in either English or Spanish.

Graduate research assistant, Ms. Sara Gonzales-Perez, a native Spanish speaker and a Master's student in Zoology handled all the Spanish references in the project.

The references were organized in alphabetic order by the last name of the first authors, and the work began with the publications by authors' names beginning with the letter "a" and proceeded until the target of 400 references was reached. The organization was convenient but also reduced biases towards geographic regions or species in the studies across the Gulf of Mexico. Dr. Simons provided references in either hard copies or digital forms. Students divided up these references for each alphabet group, and read through the references individually.

Prior to starting the coding efforts, the OU research team studied the CMECS III document and discussed the CMECS standards and coding scheme to get familiarized with the information needed for the classification system. Students were also given the following directions:

- Identify ecological and habitat information provided in the references but not included in the current CMECS III classification specifications;
- Suggest modifications to improve the CMECS III system, such as refinements to the categories or subclasses, additional codes to improve clarification, and others;
- Consider coding uncertainty information and test the idea of evidence-based uncertainty information instead of "ranking" uncertainty

Our research methodology was fundamentally empirical as with most test-based studies. The empirical approach compliments the developments of CMECS based on experts' knowledge. Students read through each reference and extracted ecological information to populate the CMECS coding table. When issues arose, we met to discuss the associated problems and proposed potential solutions. Students then tested the solutions to extract information and populated the CMECS coding table. A solution was considered working if all three students could apply the solutions to encode information in questions from different references. Otherwise, the research team met again to discuss alternatives.

Dr. Simons visited the OU team on September 25, 2009 and provided valuable expert knowledge to the team, particularly in two areas: (1) knowledge of fish species and species' scientific names; and (2) information about the CMECS working groups and issues under considerations (such as encoding temporal information). Dr. Simons continued providing references and consultation to the OU team through email and phone. He visited the team again on April 26-27, 2010 to review and approve the final deliverables before the end of the project, including the GIS database, the metadata and a number of map examples.

Results and Deliverables

The project started three months later than originally planned because of a series of unexpected administrative challenges to establish the contract. In addition, that the project workload exceeded the budgeted human resources (i.e. two student assistants) soon became apparent. We were fortunate to have an additional student assistant, Derek Morris, to work on the project with a student research award from Oklahoma NASA Space Grant for September-December 2009. The three students work diligently to catch up with the project schedule.

The project ended on May 15, 2010. The students completed coded references A-M for a total of 449 references in English or Spanish, exceeding the planned 400 references in the proposal. Each article could consist of multiple entries in the CMECS table since a study might include multiple habitat sites. Students extracted as much ecological information from the reference as possible, especially the information included in the CMECS. However, many CMECS items might not be available from a reference, and these items were left blank in the CMECS table.

All four tasks of the project were accomplished:

1. Recorded the habitat types and locations for 449 references (target was set at 400 references) on selected fish habitat studies.
2. Developed procedures to extract species and habitat information and a coding scheme to match the Coastal and Marine Ecological Classification System (CMECS). The procedures include four steps:
    a. Dr. James Simons provided scanned copies of selected references.
    b. Students read through the references andxtracted habitat information according to the CMECS specifications
    c. Issues were discussed at group meetings and alternative solutions developed. These issues were highlighted in monthly reports and are included in the issues and suggestions below.
    d. Implemented the revised codes in the geodatabase and documented meta-data
3. Built a database to record the information about species and habitats compatible with CMECS as well as information about uncertainty associated with each data item:
    a. To account for many to many relationships between fish species and habitat sites, we have developed a three-table structure: one for fish species, one for habitat information (i.e. CMECS coding table), and the other GIS feature class for locations associated with habitat information.
    b. The three tables are related through the "Site ID" to assure linkages of fish species and corresponding habitat sites. The GIS feature class also uses the Site IDs to link fish species and habitat sites to geospatial locations.

c. In preparation for developing a geodatabase for the habitats, species, and their geospatial references, we modified the field names in the CMECS table adding the underscore character "_" to accommodate database technical requirements.

d. We designed an evidence-based uncertainty coding scheme, which indicates how each value is determined during CMECS coding. Using this scheme, students identified level of uncertainty in their CMECS classification codes. The evidence-based approach provides an objective expression of how the information is acquired and leaves it to the user of the data to determine the level of confidence in the data:

   i. Explicit and direct link: information within the study was explicitly stated.

   ii. Inference from text: information within the study could be inferred but was not explicitly stated.

   iii. Extrapolated from other studies: information was extrapolated from another study.

   iv. Geographical association: information within the study is associated with its location.

   v. Educated guess: information based on the knowledge of the person coding the study.

4. Georeferenced the data set and mapped the locations of species and habitats in the Gulf of Mexico based on the data set; developed a case study (see figures below) to demonstrate meta analyses of the habitat, species studied, and location.

   Both point feature classes and polygon feature classes were created to geospatially reference habitat sites. Some studies provide location information as x and y coordinates (mostly in latitude and longitude), and other studies include maps to show the study area. Point feature classes were created to show locations based on given x and y coordinates. For study areas, maps were transcribed manually through references to existing GIS data to digitize boundaries of the study areas and create polygon feature classes, or through applications of feature matching, rectification, and transformation functions in ArcGIS to digitize these regions to polygon feature classes. Figure 1 displays all the study sites extracted from the 449 references used in the project. There are 4092 point-based study sites, 449 study areas (white polygons), and 3 survey lines (in light purple color) that were extracted from the references. Many study areas overlap, providing opportunities for future comparative studies.

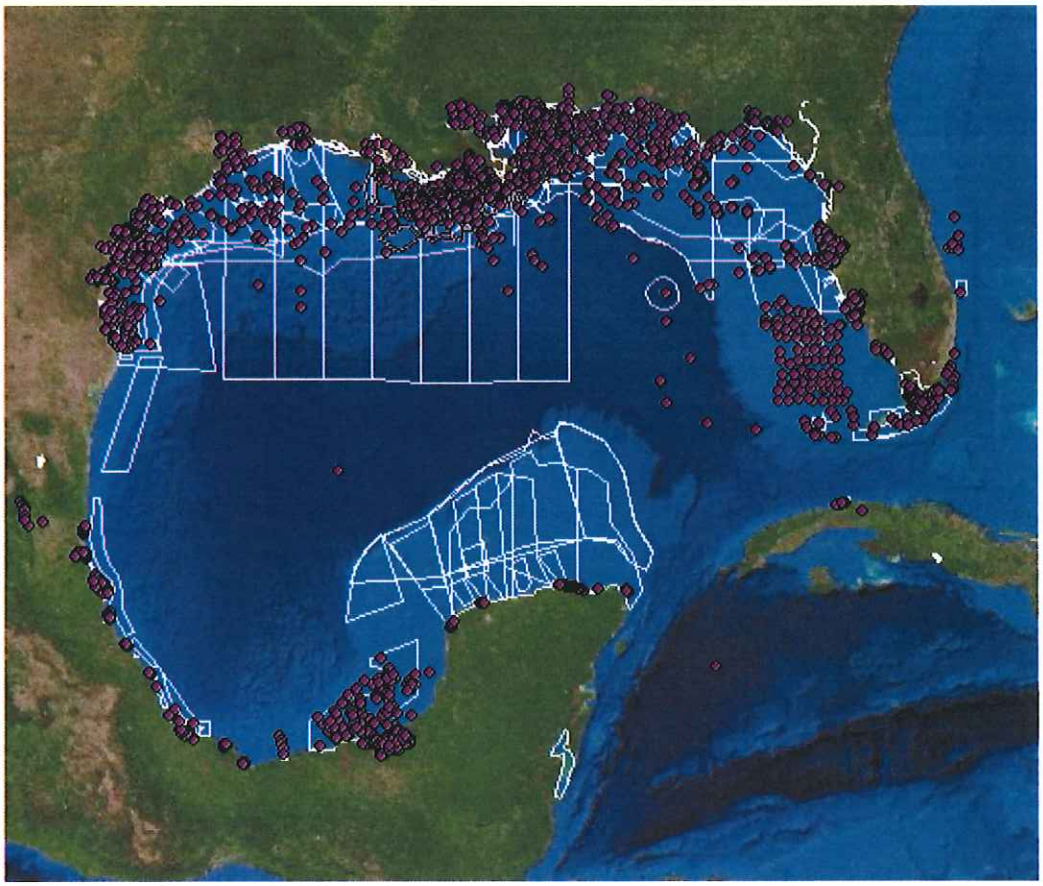
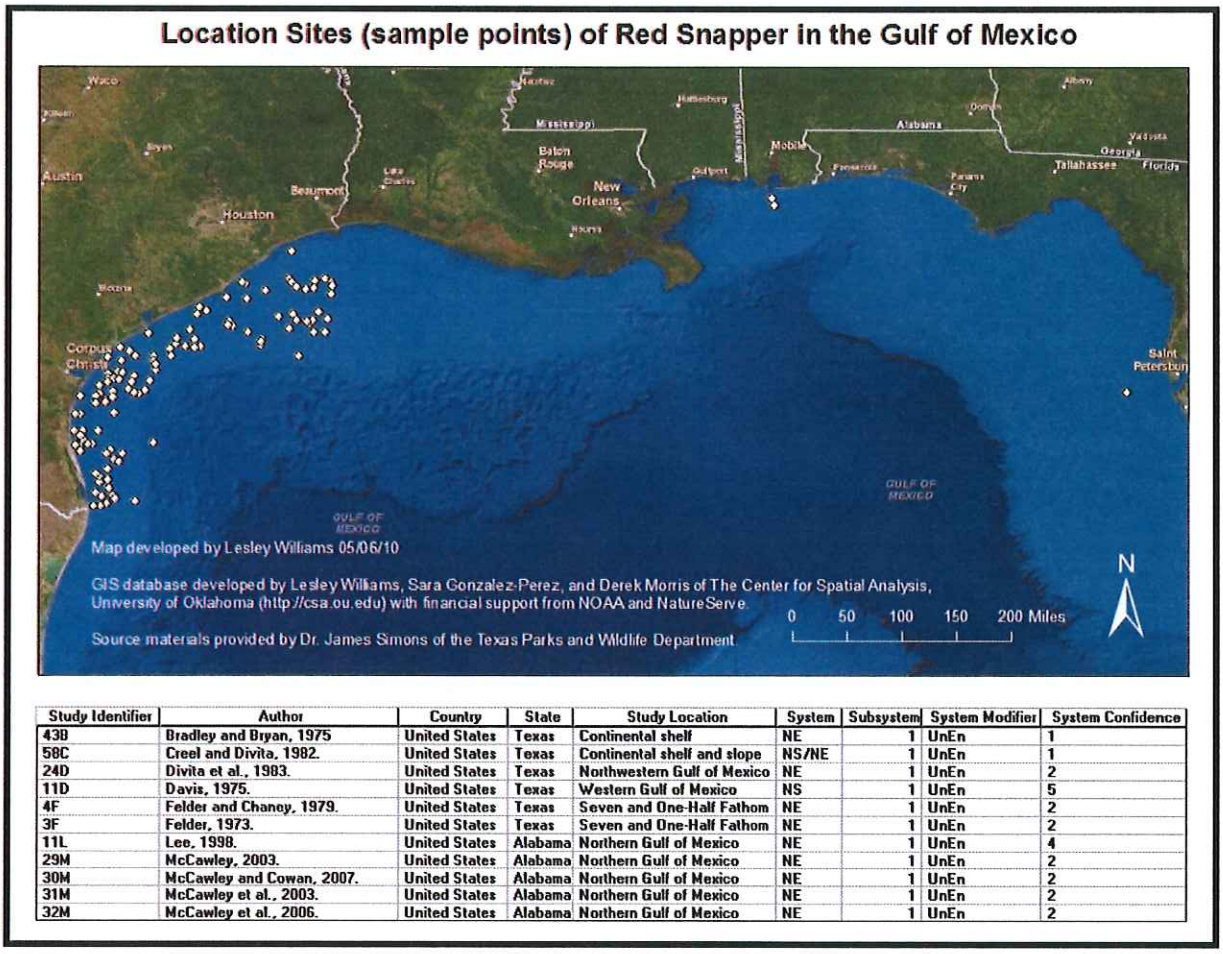


Figure 1: Study sites in the selected 449 references

The ecological habitat data extracted from the studies and coded to CMECS types, along with the study locations and the fish species are integrated in a geodatabase. Hence, the geodatabase includes habitat data, study sites (points, lines, or polygons), and fish species. The geodatabase connects these data records through study identifiers to enable three major types of queries:

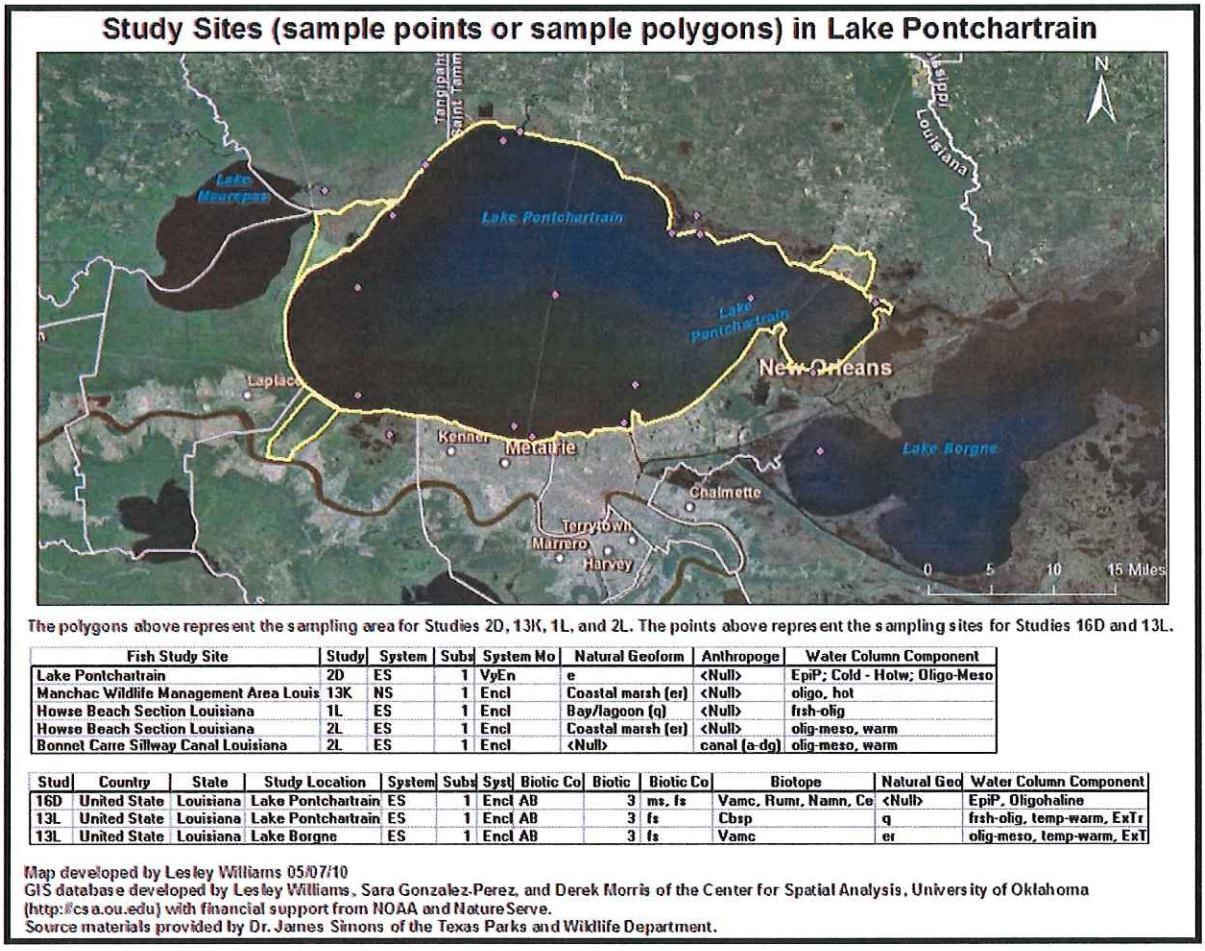
a. *Habitat queries:* What are habitats associated with the specified fish species or area of interest (example: Red Snapper in Figure 2)? The query started with specifying red snapper as the fish of interest. The figure shows study locations where red snapper were sampled. Once the study locations were selected from the geodatabase, the habitat table opened to show the associated ecological information (note: the habitat table is too long to show all fields on the page. Additional habitat data included in the table are illustrated in Figure 3).



Study Identifier	Author	Country	State	Study Location	System	Subsystem	System Modifier	System Confidence
43B	Bradley and Bryan, 1975	United States	Texas	Continental shelf	NE	1	UnEn	1
58C	Creel and Divita, 1982.	United States	Texas	Continental shelf and slope	NS/NE	1	UnEn	1
24D	Divita et al., 1983.	United States	Texas	Northwestern Gulf of Mexico	NE	1	UnEn	2
11D	Davis, 1975.	United States	Texas	Western Gulf of Mexico	NS	1	UnEn	5
4F	Felder and Chaney, 1979.	United States	Texas	Seven and One-Half Fathom	NE	1	UnEn	2
3F	Felder, 1973.	United States	Texas	Seven and One-Half Fathom	NE	1	UnEn	2
11L	Lee, 1998.	United States	Alabama	Northern Gulf of Mexico	NE	1	UnEn	4
29M	McCawley, 2003.	United States	Alabama	Northern Gulf of Mexico	NE	1	UnEn	2
30M	McCawley and Cowan, 2007.	United States	Alabama	Northern Gulf of Mexico	NE	1	UnEn	2
31M	McCawley et al., 2003.	United States	Alabama	Northern Gulf of Mexico	NE	1	UnEn	2
32M	McCawley et al., 2006.	United States	Alabama	Northern Gulf of Mexico	NE	1	UnEn	2

Figure 2: Locations of red snapper studies and habitat information

b. *Species queries:* What are species associated with the specified habitats or area of interest (Example: What are fish species being studied in Lake Pontchartrain in Figure 3)? The query process started with identification of Lake Pontchartrain. Study sites surrounding the lake show to indicate locations where fish species had been studied. Once the sites were selected, the fish species table and habitat table were opened to show the fish species found at these sites as well as associated ecological information. The habitat table is too extensive to fit on one page. Additional attribute information in other figures can be found in the other figures in these cases.



The polygons above represent the sampling area for Studies 2D, 13K, 1L, and 2L. The points above represent the sampling sites for Studies 16D and 13L.

Fish Study Site	Study	System	Subs	System Mo	Natural Geoform	Anthropoge	Water Column Component
Lake Pontchartrain	2D	ES	1	VyEn	e	<Null>	EpiP; Cold - Hotw; Oligo-Meso
Manchac Wildlife Management Area Louis	13K	NS	1	Encl	Coastal marsh (er)	<Null>	oligo, hot
Howse Beach Section Louisiana	1L	ES	1	Encl	Bay/lagoon (q)	<Null>	frsh-olig
Howse Beach Section Louisiana	2L	ES	1	Encl	Coastal marsh (er)	<Null>	olig-meso, warm
Bonnet Carre Sillway Canal Louisiana	2L	ES	1	Encl	<Null>	canal (a-dg)	olig-meso, warm

Stud	Country	State	Study Location	System	Subs	Syst	Biotic Co	Biotic	Biotic Co	Biotope	Natural Geo	Water Column Component
16D	United State	Louisiana	Lake Pontchartrain	ES	1	Encl	AB	3	ms, fs	Vamc, Rumi, Namn, Ce	<Null>	EpiP, Oligohaline
13L	United State	Louisiana	Lake Pontchartrain	ES	1	Encl	AB	3	fs	Cbsp	q	frsh-olig, temp-warm, ExTr
13L	United State	Louisiana	Lake Borgne	ES	1	Encl	AB	3	fs	Vamc	er	olig-meso, temp-warm, ExT

Map developed by Lesley Williams 05/07/10
GIS database developed by Lesley Williams, Sara Gonzalez-Perez, and Derek Morris of the Center for Spatial Analysis, University of Oklahoma (http://csa.ou.edu) with financial support from NOAA and NatureServe.
Source materials provided by Dr. James Simons of the Texas Parks and Wildlife Department.

Figure 3: Fish species study sites in Lake Pontchartrain

c. *Study queries:* What are studies associated with the specified fish species, area of interest, habitat characteristics, or any combination (Example, find all studies about Spanish mackerel in Estuarine Environment in Figure 4)? The query started with specifications of the type of fish (Spanish mackerel) and habitat of interest (estuarine environment). Sites met the criteria are selected and the corresponding entries in the habitat table and fish species table are also selected and displayed.

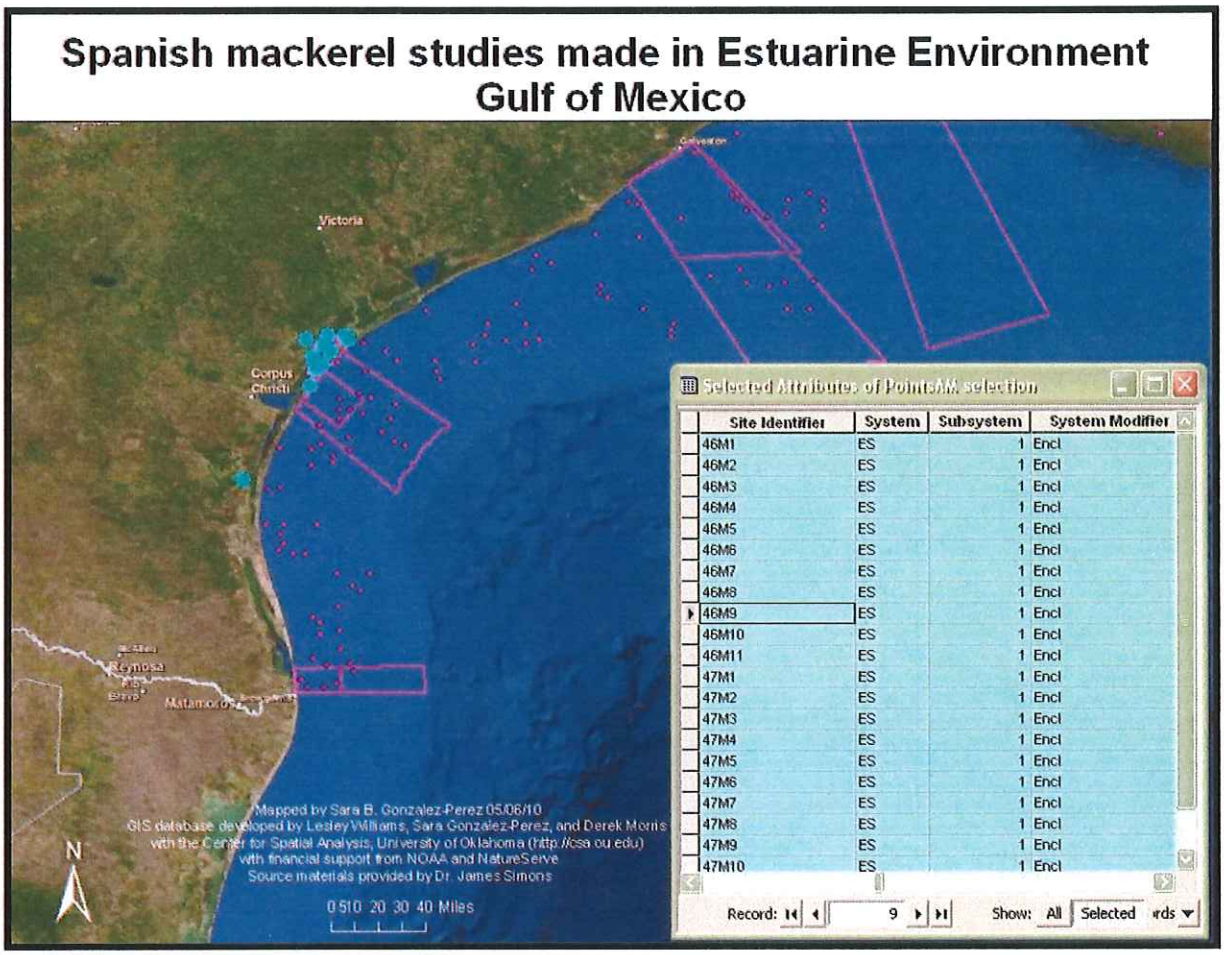


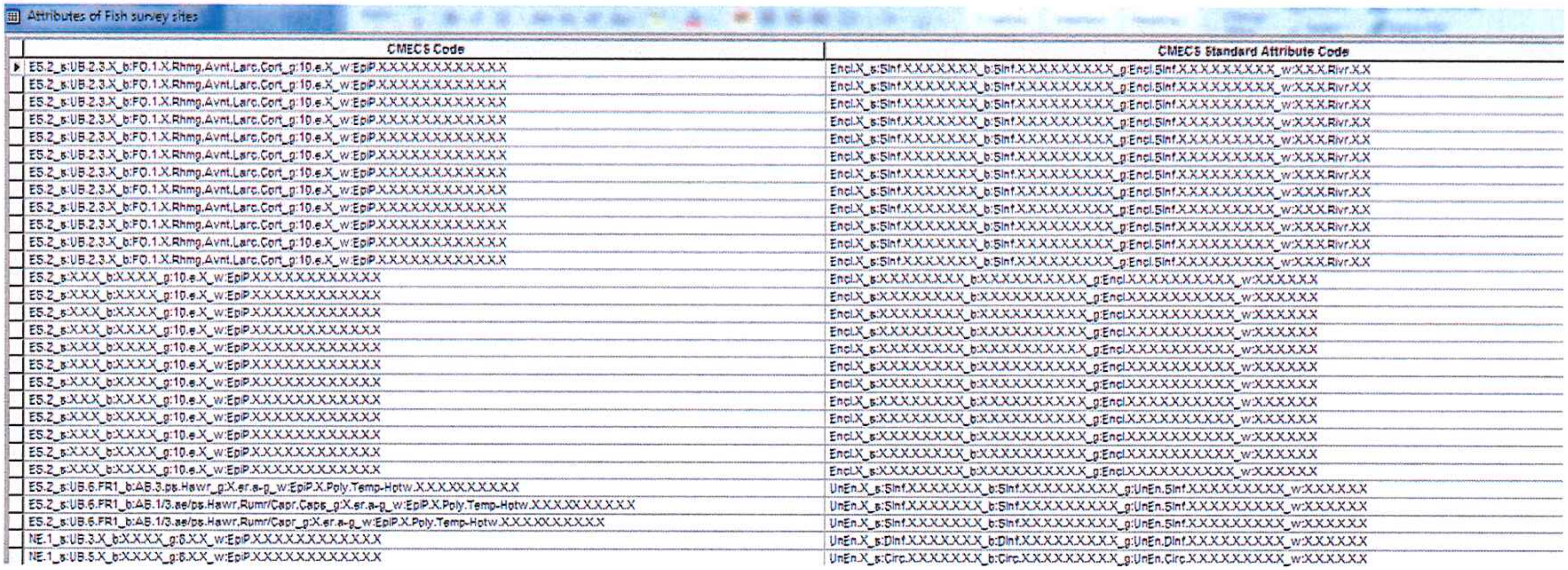
Figure 4: Spanish mackerel studies in estuarine environment in Gulf of Mexico

Issues and Suggestions

The study yielded a number of suggestions related to CMECS code structure and metadata.

1. Add temporal information to the CMECS table: published date and sample date. Additional information may be necessary, such that salinity changes seasonally (the four seasons or dry-wet season, etc). Issues related to temporal changes in coastal marine environments are being considered by a CMECS working group; hence, we did not attempt to resolve these in this study.
2. The last column of the CMECS coding table is for CMECS code structure. Examples provided in the CMECS document (p. 71-73) suggest leaving off the end of the string if the data at the end of the string are missing and inserting "0" or "00" when missing data are in the middle of the string. This results in strings of different lengths and may create challenges in data queries or programming. In addition, use of "0" for missing data can create confusion between no data and data of value "0". We suggest the following to address the issue:
    a. Make all codes the same length.
    b. For missing data regardless of where the missing data are located in the code structure, use "X" instead of "0".
    c. For data that cannot fit into the existing classification items of categories or subcategories, use "?".

    The current coding scheme in the developed geodatabase follows the above suggestion. **Error! Reference source not found.** is an example of how the codes appear in the geodatabase.


Attributes of Fish survey sites

CMECS Code	CMECS Standard Attribute Code
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:UB.2.3.X_b:FO.1.X.Rhmg.Avnt.Larc.Cort_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:Encl.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.Rivr.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:X.X.X_b:X.X.X.X_g:10.e.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	Encl.X_s:X.X.X.X.X.X.X.X_b:X.X.X.X.X.X.X.X.X.X_g:Encl.X.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:UB.6.FR1_b:AB.3.ps.Hawr_g:X.er.a-g_w:EpiP.X.Poly.Temp-Hotw.X.X.X.XX.X.X.X.X.X	UnEn.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:UnEn.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:UB.6.FR1_b:AB.1/3.ae/ps.Hawr,Rumr/Capr,Caps_g:X.er.a-g_w:EpiP.X.Poly.Temp-Hotw.X.X.X.XX.X.X.X.X.X	UnEn.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:UnEn.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
E5.2_s:UB.6.FR1_b:AB.1/3.ae/ps.Hawr,Rumr/Capr_g:X.er.a-g_w:EpiP.X.Poly.Temp-Hotw.X.X.X.XX.X.X.X.X.X	UnEn.X_s:5Inf.X.X.X.X.X.X.X_b:5Inf.X.X.X.X.X.X.X.X.X_g:UnEn.5Inf.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
NE.1_s:UB.3.X_b:X.X.X.X_g:8.X.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	UnEn.X_s:DInf.X.X.X.X.X.X.X_b:DInf.X.X.X.X.X.X.X.X.X_g:UnEn.DInf.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X
NE.1_s:UB.5.X_b:X.X.X.X_g:8.X.X_w:EpiP.X.X.X.X.X.X.X.X.X.X.X.X.X	UnEn.X_s:Circ.X.X.X.X.X.X.X_b:Circ.X.X.X.X.X.X.X.X.X_g:UnEn.Circ.X.X.X.X.X.X.X.X.X_w:X.X.X.X.X.X

Figure 5: Examples of CMECS codes and CMECS standard attribute codes. The CMECS_Code covers System, Subsystem, Surface Geology Component, etc. The CMECS standard attribute coded are the classifiers and modifiers (Standard Attributes) that further describe CMECS types.

3. Metadata documentation

    a. We compiled metadata for the CMECS and the species data tables. Each field name (column in a table) and the potential field values for the field are specified.

    b. The metadata are compiled in ArcGIS and are in compliance with FGDC metadata standards. Hence, the metadata are readily applicable to the geospatial database converted from the CMECS table and species table with geospatial references as one of the final deliverables for the project.

4. Identified coding issues

    a. Many studies do not provide detailed descriptions on habitat environments, and many of them have no information at all. Therefore, many entries in the CMECS scheme are empty

    b. For certain study sites it is difficult to make inferences about habitat information. One study refers to a reef named One-Half Fathom Reef but provides no information about the reef. At the other extreme is another example which referred to "well indurated calichified lake marl." Such specialized terminology was not always interpretable by the students working on this study.

    c. A broader issue is that the CMECS scheme may need to have another entry for studies that only provide broad categories of data for certain CMECS modifiers or classifiers. Some studies have precise information that matches well with CMECS categories or classes, but others only have general information. For example, some studies give average temperature, and others only provide temperature gradients. Turbidity is another example. The CMECS coding scheme references turbidity by measurements under the Secchi Depth Reading. Some studies do not provide the Secchi Depth Readings or any kind of measurements. They describe turbidity as usually turbid or highly turbid.

    d. Some species attach themselves to other species, such as sharksuckers to dolphins. Should we consider the habitat of dolphins also the habitat for sharksuckers? Should the CMECS scheme add an entry to document such cases?

    e. CMECS only considers surface waves generated by wind or seismic energy. Some studies suggest habitat sites with waves generated by boating or human activities.

    f. CMECS has no subclass for volcanic ashes which we think should be added to the Unconsolidated Bottom class.

    g. CMECS does not consider "situational" habitat. Some studies indicate that vegetation varies along shorelines and would become part of a habitat only during extreme high tides. Some studies indicate pools formed during receding hurricane waters and

temporary rainwater pools. Moreover, some studies indicate that there is mud in the center of a habitat with sand around the edges. CMECS only has a subclass for mixed sediments but in this case, the mud and sand are not mixed.

h. The following are additional cases from the studies where it was difficult to identify the proper coding strategy:
    i. The bottom smells of hydrogen sulfide
    ii. Artesian well as being a constant supply to the center of a lagoon
    iii. Man-made water sources
    iv. Rain quickly reducing salinity in some areas of a habitat; salinity increases due to evaporation in hot weather
    v. One-Half Fathom Reef, only referenced by name, but provides no other habitat information

Conclusions

In conclusion, this study has compiled data on fish habitats and locations throughout the Gulf of Mexico from 449 studies into a geodatabase, providing both valuable information about the ecology of the Gulf and important input into the Coastal and Marine Ecological Classification Standard (CMECS). We believe that this information is particularly relevant now that the Gulf is experiencing extensive environmental impact from the current massive oil spill. We hope that it will prove helpful in guiding assessment and mitigation of damages. In addition, it should be useful in improving and developing mapping guidance for CMECS.

Appendix: Meta Data Explanation

All metadata are associated with the geodatabase. The list below is provided for additional reference. CMECS Metadata is listed following each attribute heading. Within the properties of each table or feature class, an alias can be assigned to better understand the attribute's heading. Each alias is listed below following the metadata information.

Purpose:

Coastal and marine planners and managers are faced with a complex environment in which to make difficult decisions about habitat conservation and resource management. Given this complexity and recognizing the vastness of the marine environment for which the United States has jurisdiction, there is an urgent and increasing need for a national habitat classification approach that can be used to develop strategies for resource management and conservation. To meet this need NOAA and NatureServe developed the Coastal and Marine Ecological Classification Standard (CMECS), a classification standard that is relevant to all U.S. coastal and marine environments and that can be applied on local, regional and continental scales. The continuing goal of the CMECS standard is to permit the classification of ecological and habitat units within a simple standard format that uses a common terminology. This is intended to allow effective identification, monitoring, protection, and restoration of unique biotic assemblages, protected species, critical habitat, and important ecosystem components. This classification standard will provide a uniform protocol for identifying and naming both new and existing ecological units.

Abstract:

CMECS Version III first classifies the coastal and marine environment to describe an aquatic setting, then provides detail through five distinct components that each describe a different aspect of the relevant ecology. These components provide a structured way to organize information about coastal and marine habitats, and provide a standard terminology for describing them. The aquatic setting is identified as a System in broad terms: Nearshore, Neritic, Oceanic, Estuarine, Freshwater-Influenced, or Lacustrine. The Surface Geology Component (SGC) then describes the geological composition and environment of the upper few centimeters of substrate, describing also the structural aspects of biogenic substrates and settings, such as coral reefs. The Biotic Cover Component (BCC) is a hierarchical classification that describes the biological composition and cover of the coastal and marine benthos [example types include Clam Bed, Spur and Groove Reef, and Rooted Vascular Vegetation]. The BCC and the SGC are designed to be used together to more thoroughly describe the benthic environment and to describe benthic cover as seen from above. The Sub-Benthic Component (SBC) describes characteristics of the sediments and soils with depth, providing more detailed information on the composition of the entire sediment column. The Geoform Component (GFC) describes the major geomorphic or structural characteristics of the coast and seafloor at various scales (example units include abyssal plain, seamount, delta, and beach). The Water Column Component (WCC) describes the structure, patterns and processes of the overlying water column (example types are defined by depth, water column structure, hydrologic formation, salinity, temperature and biotic composition). Each of the five components describes distinct conceptual ecological units. Each component is intended to be identified and mapped independently or combined as needed to inform specific questions or problems. In addition to the five components, CMECS provides a list of standard attributes -- a consistent set of variables that provide the basis for classification

and description of the CMECS units. When required to define a unit, these standard attributes are called "classifiers". When not required for classification, but used to further describe a unit, these standard attributes are called "modifiers". Modifiers provide users the ability to customize their application of the classification in a standardized way.

Entity Attribute:

CMECS_A_M: Data table of the CMECS coded studies whose authors last names begin with the letters A through M.

Study_ID: Each study has been given a number which includes the first letter of the author's last name. Alias: Study Identifier

Pub_Date: The published date of the journal article, thesis, or abstract. Alias: Published Date

Country: The country where the sampling of fish species occurred. Alias: Country

State: The state nearest to which the sampling of fish species occurred. Alias: State

Location: The country, state, city, and/or name of the study site where sampling occurred. Alias: Location

Study_Location: The specific name of the location where sampling of fish species occurred. Alias: Study Location

Sta_Lat: The latitude of where the sampling of fish species occurred. Alias: Station Latitude

Sta_Long: The longitude of where the sampling of fish species occurred. Alias: Station Longitude

Month: The month in which the sampling of fish species occurred. Alias: Month

Year: The year in which the sampling of fish species occurred. Alias: Year

FishAM: Polygon feature class indicating location site where fish were caught for sampling.

Points_A_M: Data table indicating location site where fish were caught for sampling.

PointsAM: Point feature class indicating location site where fish were caught for sampling.

fishstudy: Name of the location site where fish were caught for sampling. Alias: Fish Study Site

Author: The author's last name of the study. Alias: Author

Lines_24F: Polyline feature class indicating the location where fish were caught for sampling.

Site_ID: An identifier that relates the CMECS coded table to the Species table and to the feature classes. Alias: Site Identifier

Spec_Name: The scientific name of each fish species. Alias: Species Name

Species_A_M: Data table of the study identifier, species name, and the site identifier.

System: Systems are differentiated from one another by a combination of salinity, geomorphology and depth. Salinity is first used to separate the truly marine systems from those influenced by freshwater. Three systems, Nearshore [NS], Neritic [NE], and Oceanic [OC], are truly marine -- all having salinities greater than 30 PSU throughout the year. They are distinguished from each other by depth and relative distance from the continental shelf. Two Systems, Estuarine [ES], and Freshwater-Influenced [FI], are at least occasionally diluted (< 30PSU) by significant freshwater input during the year. The Estuarine and Freshwater-Influenced Systems are distinguished from each other by the degree of enclosure by land. The Lacustrine System [LA]is used to classify the U.S. Great Lakes and is defined as having salinity less than 0.5 PSU. Alias: System

Subsystem: **Subtidal [1] –** The substrate is continuously submerged relative to the extreme low tide line. **Intertidal [2] –** The substrate is regularly and periodically exposed and flooded by tides. This zone includes the supratidal zone -- the area above the high tide line in the splash zone that is affected by spray, splash, aerosols and overwash. Limnetic [1] – all deepwater habitats within the Lacustrine System greater than 2 m. Littoral [2] - All wetland habitats in the Lacustrine System. Extends from the shoreward boundary of the system to a depth of 2 m below low water or to the maximum extent of nonpersistent emergents, if these grow at depths greater than 2 m. Alias: Subsystem

System_Mod: Enclosure represents the degree of isolation of a water body from other waters due to enclosure by a land mass. Unenclosed [UnEn]: > 150° angular gap from landward end of water body to seaward opening. No confining land masses (*e.g.*, islands) within or just outside water body. Partially enclosed [PrEn]: 90° - 150° angular gap from landward end of water body to seaward opening. Significantly enclosed [SgEn]: 45° - 90° angular gap from landward end of water body to seaward opening. Very enclosed [VyEn]: 10° - 45° angular gap from landward end of water body to seaward opening. Enclosed [Encl]: Essentially separated from the ocean. Intermittent [Intm]: Class of water bodies that regularly close due to low flow, opening seasonally during high flow. Position Relative to Shelf Break is a classifier indicating the position of an area of the ocean relative to the edge of the Continental Shelf. Seaward of the shelf break [Seaw] - Where the continental platform begins to steepen toward the continental slope, defines the Oceanic System. Landward of the shelf break [Lndw] – defines the outer limit Neritic System. Alias: System Modifier

Sys_Conf: The level of confidence in which a study was coded. Each column has been given a number indicating the confidence level. The number 1 is explicit and direct link, number 2 is inference from text, number 3 is extrapolated from other studies, number 4 is geographical associations, and number 5 is educated guess. Explicit and direct link: information within the study was explicitly stated. Inference from text: information within the study could be inferred but was not explicitly stated. Extrapolated from other studies: information was extrapolated from another study. Geographical association: information within the study is associated with its location. Educated guess: information based on the knowledge of the person coding the study. Alias: System Confidence

Surf_Geo_Com, Surf_Geo_Sub, and Reef_Mor: Classes for the SGC are determined by the dominant geologic or biogenic cover of the substrate. Subclasses are defined by the composition and particle size of the substrate, or, in the case of biogenic reef substrates, by reef geomorphology. Class: Rock Bottom [RB] – Subclass: Bedrock [1], Boulder [2], Pavement [3]

Class: Unconsolidated Bottom: [UB] – Subclass: Cobble/gravel [1], Sands [2], Mud [3], Organic [4], Shell [5], Mixed Sediments [6] Class: Rocky Shore [RS] – Subclass: Bedrock [1], Boulder [2], Pavement [3] Class: Unconsolidated Shore [US] – Subclass: Cobble/gravel [1], Sands [2], Mud [3], Organic [4], Shell [5], Mixed Sediments [6] Class: Coral Reef [CR] – Subclasses: Reef Lagoon [1], Back Reef [2], Reef Flat [3], Reef Crest [4], Forereef [5], Deep Forereef [6], Pinnacle Reef [7], Mesophotic Reef [8], Deep/Cold Water Reef [9], Outlier Reef [10] – Reef Morphology: Spur and Groove Reef [a], Patch Reef [b], Aggregate Patch [c], Linear Reef [d], Aggregate Reef [e], Live Hardbottom [f], Scattered Coral/Rock [g], Unconsolidated Sands [h], Unconsolidated Muds [i] Class: Faunal Reef [FR] – Subclass: Mollusk Reef [1], Worm Reef [2]. Alias: Surface Geology Component, Surface Geology Subclass, and Reef Morphology

Surf_Mod: Benthic Depth Zones: Littoral [Litt]: intertidal, Shallow Infralittoral [Sinf]: 0-5 m, Deep Infralittoral [Dinf]: 5-30 m, Circalittoral [Circ]: 30-200 m, Mesobenthic [MsoB]: 200-1000 m, Bathybenthic [BtyB]: 1000-3500 m, Abyssalbenthic [AbyB]: 3500-6000 m, Hadalbenthic [HadB]: >6000 m. Energy Intesity: Very low energy [VLEn], Low energy [LoEn], Moderate energy [MoEn], High energy [HiEn]. Tide Range: Microtidal [MiTd]: <0.1m, Small tide range [SmTd]: 0.1-1m, Moderate tide range [MoTd]: 1-5m, Large tide range [LgTd]:>5m. Profile: None [NoPr]: 0, Low [LoPr]: 0-2m, Medium [MdPr]: 2-5m, High [HiPr]: >5m. Slope: Flat [Flat] 0-5°, Sloping [Slpg]: 5-30°, Steeply sloping [SSlp]: 30-45°, Vertical [Vert]: 45-90°, Overhang [Ohng]:>90°. Anthropogenic Impact: Developed, Impounded/Diverted [Impd], Dredged area/Channel [Drdg], Filled [Flld], Contaminated [Cont]. Rugosity: Very Low [VloR]: 1.0-1.25, Low [LowR]: 1.25-1.50, Moderate [ModR]: 1.50-1.75, High [HghR]: 1.75-2.00, Very High [VhiR]: >2.00. Temporal Persistence: Stochastic [Stoc], Hours [Hors], Days [Days], Weeks [Weks], Months [Mnts], Seasons [Seas], Years [Yers], Inter-annual [IntA], Decades [Decs], Centuries [Cent]. Alias: Surface Modifier

Surf_Conf: The level of confidence in which a study was coded. Each column has been given a number indicating the confidence level. The number 1 is explicit and direct link, number 2 is inference from text, number 3 is extrapolated from other studies, number 4 is geographical associations, and number 5 is educated guess. Explicit and direct link: information within the study was explicitly stated. Inference from text: information within the study could be inferred but was not explicitly stated. Extrapolated from other studies: information was extrapolated from another study. Geographical association: information within the study is associated with its location. Educated guess: information based on the knowledge of the person coding the study. Alias: Surface Confidence

Bio_Cover: The biological component of CMECS is a classification of the biotic aspects of substrate at different spatial scales, and is intended to be used together with the SGC so that biology is considered in the context of physical habitat. Class: Faunal Reef [FR], Coral Reef [CR], Faunal Bed [FB], Aquatic Bed [AB], Emergent/Low Shrub Wetland [EM], Forested Wetland [FO]. Alias: Biotic Cover Component

Bio_Sub: Class: Faunal Reef [FR] – Subclass: Mollusk Reef [1], Worm Reef [2], Class: Coral Reef [CR] – Subclasses: Living Stony Coral Communities [1], Calcareous Algal Communities [2], Class: Faunal Bed [FB] – Subclasses: Sessile Epifauna [1], Mobile Fauna [2], Infauna [3], Class: Aquatic Bed [AB] – Subclasses: Macroalgae [1], Rooted Vascular [3], Microbial Mat [5?] , Class: Emergent/Low Shrub Wetland [EM] – Subclasses: Coastal Marsh [1], Tidal Shrubland [2], Class: Forested Wetland [FO] – Subclass: Mangrove [1]. Alias: Biotic Cover Subclass

Bio_Group: Class: Faunal Reef [FR] – Subclass: Mollusk Reef [1] – Biotic Group: Oyster Reef [or], Mussel Reef [mr], Gastropod Reef [gr]; Subclass: Worm Reef [2]- Biotic Group: Sabellariid Reef [sr]; Class: Coral Reef [CR] – Subclass: Living Stony Coral Communities [1] – Biotic Group: Robust Branching Corals [bc], Fragile Branching Corals [fb], Table Corals [tc], Massive Corals [ma], Plate Corals [pc], Encrusting Corals [en]; Subclass: Calcareous Algal Communities [2] – Biotic Group: Rhodolith Beds [rd], Crustose Calcareous Algae [cc], Class: Faunal Bed [FB] – Subclass: Sessile Epifauna [1] – Biotic Group: Oyster Bed [ob], Mussel Bed [mb], Sessile Gastropods [sg], Barnacles [bn], Coral Garden [cg], Mixed Colonizers [xc], Sponge Bed [sp], Attached Anemones [aa], Burrowing Anemones [ba], Small Tube Building Fauna [st], Large Tube Building Fauna [lt], Crinoids [cn], Hydroids [hy], Bryozoans [br], Tunicate Bed [tb], Foraminifera [fm], Subclass: Mobile Fauna [2]- Biotic Group: Mobile Gastropods [mg], Mobile Crustaceans [mc], Scallop Beds [sb], Sand Dollars [sd], Ophiuroids [op], Holothurians [ho], Subclass: Infauna [3] – Biotic Group: Clam Bed [cb], Tunneling Megafauna [tm], Small Surface Burrowing Fauna [ss], Deposit Feeders [df], larger Deep Burrowing Fauna [db], Burrowing Urchins [bu], Ophiuroids [op], Echiurid Communities [ec], Oligozoic [oz], Class: Aquatic Bed [AB] – Subclass: Macroalgae [1]- Biotic Group: Attached ephemeral macroalgae [ae], Rockweeds [rw], Leathery Red Macroalgae [lr], Kelp Forests [kf], Leathery Green Macroalgae [lg], Jointed Calcareous Algae [jc], Attached Crustose Algae [ac], Rhodolith Beds [rd], Drift Ephemeral Algae [de], Drift Algae [da], Subclass: Rooted Vascular [3] – Biotic Group: Poly/Euhaline Seagrass Beds [ps], Oligo/Mesohaline Seagrass Beds [ms], Fresh/Oligohaline Seagrass Beds [fs], Subclass: Microbial Mat [5?]- Biotic Group: Microphytobenthos [mp], Bacterial mat [bm], Chemautotrophic Bacteria [ca] , Class: Emergent/Low Shrub Wetland [EM] – Subclass: Coastal Marsh [1]- Biotic Group: High Marsh [hm], Low Marsh [lm]. Alias: Biotic Cover Group

Biotope: Scientific names of plants identified within each study. Alias: Biotope

Bio_Mod: Benthic Depth Zones: Littoral [Litt]: intertidal, Shallow Infralittoral [Sinf]: 0-5 m, Deep Infralittoral [Dinf]: 5-30 m, Circalittoral [Circ]: 30-200 m, Mesobenthic [MsoB]: 200-1000 m, Bathybenthic [BtyB]: 1000-3500 m, Abyssalbenthic [AbyB]: 3500-6000 m, Hadalbenthic [HadB]: >6000 m. Percent Cover Range: Bare [Bare]: <1% cover, Sparse [Sprs]: 1-10% cover; Moderately Sparse [MdSp]: 10-25% cover, Moderate [Modt]: 25-75% cover, Dense [Dens]: 75-90% cover, Complete [Cmpl]: 90-100% cover. Energy Intesity: Very low energy [VLEn], Low energy [LoEn], Moderate energy [MoEn], High energy [HiEn]. Tide Range: Microtidal [MiTd]: <0.1m, Small tide range [SmTd]: 0.1-1m, Moderate tide range [MoTd]: 1-5m, Large tide range [LgTd]:>5m. Profile: None [NoPr]: 0, Low [LoPr]: 0-2m, Medium [MdPr]: 2-5m, High [HiPr]: >5m. Primary Water Source: Watershed [Wshd], Local estuary exchange [LoEs], Local ocean exchange [LoOc], River [Rivr], Estuary [Estu], Marine [Mrin]. Slope: Flat [Flat] 0-5°, Sloping [Slpg]: 5-30°, Steeply sloping [SSlp]: 30-45°, Vertical [Vert]: 45-90°, Overhang [Ohng]:>90°. Anthropogenic Impact: Developed, Impounded/Diverted [Impd], Dredged area/Channel [Drdg], Filled [Flld], Contaminated [Cont]. Rugosity: Very Low [VloR]: 1.0-1.25, Low [LowR]: 1.25-1.50, Moderate [ModR]: 1.50-1.75, High [HghR]: 1.75-2.00, Very High [VhiR]: >2.00. Temporal Persistence: Stochastic [Stoc], Hours [Hors], Days [Days], Weeks [Weks], Months [Mnts], Seasons [Seas], Years [Yers], Inter-annual [IntA], Decades [Decs], Centuries [Cent]. Alias: Biotic Cover Modifier

Bio_Conf: The level of confidence in which a study was coded. Each column has been given a number indicating the confidence level. The number 1 is explicit and direct link, number 2 is inference from text, number 3 is extrapolated from other studies, number 4 is geographical associations, and number 5 is educated guess. Explicit and direct link: information within the study was explicitly stated. Inference from text: information within the study could be inferred but was not explicitly stated. Extrapolated from other studies: information was extrapolated from another study. Geographical association: information within the study is associated with its location. Educated guess: information based on the knowledge of the person coding the study. Alias: Biotic Cover Component

Sub_Ben_Com: The specifics of this component are currently being further developed by the Mapping Partnership for Coastal Soils and Sediment. Alias: Sub-Benthic Component

Physio_Prov: The Physiographic Province describes the major components of the seafloor geomorphology along the continuum from the spreading center to the coast. Fracture zone, spreading center [1], Mid-ocean ridge [2], Abyssal plain [3], Oceanic bank [4], Continental, Island rise [5], Continental, Island slope [6], Shelf break [7], Continental shelf, island shelf [8], Basin floor, borderland [10], Inland sea, enclosed sea [11]. Alias: Physiographic Province

Nat_Geo: Geoforms are seafloor structures. Apron, Deep fan, Bajada [A]; Atoll [a]; Bank [m/f]; Basin [h]; Bay, Embayment, Sound, Bight, Fjord [q]; Beach, relic (submerged) [b]; Boulder(s) [h(b)]; Canyon [c]; Canyon head [c(h)]; Canyon mouth [c(m)]; Channel, Gully, Inlet, Tidal channel [g]; Channel bank [g/m]; Delta, Fan [y]; Depression [h]; Face [_i]; Riverine Estuary [er]; Rubble zone [h(b)l_h]; Sand ripple [_r]; Scarp, Cliff, Fault, Slump scar [s]; Seamount [x]; Seamount crown, crest, top [x(c)]; Seamount base [x(b)]; Guyot, Flat-topped seamount [x(f)]; Guyot base [x/f(b)]; Sediment/sand wave [w(w)]; Sediment/sand dunes [w(d)]; Shoal, sandbar [sl]; Slough [h/g]; Solution pit, sink, karst [k]; Estuary [e]; Sub-estuary [es]; Terrace, plain [t]; Terrace/Plain – volcanic [t_v]; Trench (natural) [T]; Wall [(w)]; Vent [_e]; Tidepool [u]; Flank [F]; Flat, Floor, Seabed [f]; Fracture, Crack, Crevice, Notch, Groove [_f]; Hole, Pit, Scour Mark, Pockmark (non-karst) [h_e}; Ice feature [i]; Lagoon, Enclosed water [n]; Landslide, Slump [l]; Lava field [f_v]; Ledge, Overhang [_d]; Moraine [i_m]; Mound, Ridge, Knob [m]; Overbank deposit, Levee(natural) [o]; Pinnacle, cone [p]; Rill (linear deposit or depression) [r]; Rock outcrop [er]; Beach [b]. Alias: Natural Geoform

Anthro_Geo: Common anthropogenic geoforms. Artificial reef (a-r); Berm (a-b/m); Dam, Dike (a-g); Dredge deposit/Mound (a-dm); Dredged channel, groove, trench or hole (a-dg); Drilling platform (a-s); Harbor, Marina (a-m); Jetty (a-g); Levee (a-o); Pier (a-s); Seawall (a-s/w); Shipwreck (a-w); Trawl disturbance (a-td); Scar/Prop scar (a-f); Pilings (a-s); Archaeological feaure (a). Alias: Anthropogenic Geoform

Geo_Mod: Enclosure represents the degree of isolation of a water body from other waters due to enclosure by a land mass. Unenclosed [UnEn]: > 150° angular gap from landward end of water body to seaward opening. No confining land masses (*e.g.*, islands) within or just outside water body. Partially enclosed [PrEn]: 90° - 150° angular gap from landward end of water body to seaward opening. Significantly enclosed [SgEn]: 45° - 90° angular gap from landward end of water body to seaward opening. Very enclosed [VyEn]: 10° - 45° angular gap from landward end of water body to seaward opening. Enclosed [Encl]: Essentially separated from the ocean. Benthic Depth Zones: Littoral [Litt]: intertidal, Shallow Infralittoral [Sinf]: 0-5 m, Deep

Infralittoral [Dinf]: 5-30 m, Circalittoral [Circ]: 30-200 m, Mesobenthic [MsoB]: 200-1000 m, Bathybenthic [BtyB]: 1000-3500 m, Abyssalbenthic [AbyB]: 3500-6000 m, Hadalbenthic [HadB]: >6000 m. Energy Type: Wind [Wind]; Current [Curr]; Surface wave [SWav]; Internal Wave [IWav]; Tide [Tide]. Energy Intesity: Very low energy [VLEn], Low energy [LoEn], Moderate energy [MoEn], High energy [HiEn]. Tide Range: Microtidal [MiTd]: <0.1m, Small tide range [SmTd]: 0.1-1m, Moderate tide range [MoTd]: 1-5m, Large tide range [LgTd]:>5m. Primary Water Source: Watershed [Wshd], Local estuary exchange [LoEs], Local ocean exchange [LoOc], River [Rivr], Estuary [Estu], Marine [Mrin]. Profile: None [NoPr]: 0, Low [LoPr]: 0-2m, Medium [MdPr]: 2-5m, High [HiPr]: >5m. Slope: Flat [Flat] 0-5°, Sloping [Slpg]: 5-30°, Steeply sloping [SSlp]: 30-45°, Vertical [Vert]: 45-90°, Overhang [Ohng]:>90°. Anthropogenic Impact: Developed, Impounded/Diverted [Impd], Dredged area/Channel [Drdg], Filled [Flld], Contaminated [Cont]. Rugosity: Very Low [VloR]: 1.0-1.25, Low [LowR]: 1.25-1.50, Moderate [ModR]: 1.50-1.75, High [HghR]: 1.75-2.00, Very High [VhiR]: >2.00. Temporal Persistence: Stochastic [Stoc], Hours [Hors], Days [Days], Weeks [Weks], Months [Mnts], Seasons [Seas], Years [Yers], Inter-annual [IntA], Decades [Decs], Centuries [Cent]. Alias: Geoform Modifier

Geo_Conf: The level of confidence in which a study was coded. Each column has been given a number indicating the confidence level. The number 1 is explicit and direct link, number 2 is inference from text, number 3 is extrapolated from other studies, number 4 is geographical associations, and number 5 is educated guess. Explicit and direct link: information within the study was explicitly stated. Inference from text: information within the study could be inferred but was not explicitly stated. Extrapolated from other studies: information was extrapolated from another study. Geographical association: information within the study is associated with its location. Educated guess: information based on the knowledge of the person coding the study. Alias: Geoform Confidence

Water_Col_Com: The Water Column Component (WCC) describes the structures, patterns, physical properties, processes and associated biology of the subtidal water column. Water Column Depth Zones: Sea Surface [Surf]: 0 m; Epipelagic [EpiP]: >0-200m; Mesopelagic [MsoP]: 200-1000m; Bathylpelagic [BtyP]: 1000-4000m; Abyssalpelagic [AbyP]: 3500-6000m; Hadalpelagic [HadP]: >6000m. Salinity is grouped into categories in units of PSU (practical salinity units, nearly equivalent to PPT, parts per thousand). Fresh – 0PSU; Oligohaline - >0-5PSU; Mesohaline – 5-18PSU; Polyhaline – 18-30PSU; Euhaline – 30-40PSU; Hyperhaline - >40PSU. Temperature: Frozen [Froz] $\leq 0°$ C with surface ice; Superchilled [SChd] $\leq 0°$ C without ice; Cold [Cold] 0-10°C; Temperate [Temp] 10-20°C; Warm [Warm] 20-30°C; Hot [Hotw] >30°C. Alias: Water Column Component

Water_Par: WCC Assessment Parameters: Oxygen: Anoxic [Anox] 0-0.1mg/L; Severely Hypoxic [SHpx] 0.1-2.0 mg/L; Hypoxic [Hypx] 2.1-4mg/L; Oxic [Oxic] 4-10 mg/L; Oxygen saturated [OxSt] 10-12 mg/L; Oxygen supersaturated [OxSS] >12 mg/L. Turbidity: Extremely Turbid [ExTr] <1 m; HighlyTurbid [HiTr] 1-2m; Moderately Turbid [MdTr] 2-5m; Clear [Cler] 5-20m; Extremely clear [XClr] >20m. Turbidity Types: Chlorophyll [Chlr]; Mineral particulates [Mnrl]; Carbonate particulates [Carb]; Colloidal precipitates [Coll]; Dissolved color [DCol]; Detritus [Detr]; Mixed [MxTr]. Turbidity Provenance: Autochthonous [Auto]; Allochthonous [Allo]; Resuspended [RSsp]; Precipitated [Prcp]; Terrigenous origin [Terr]; Marine origin [Mari]; Photic Quality: Photic [Phot]; Dysphotic [Dysp]; Aphotic [Apho]. Productivity:

Productivity is a general categorization of the abundance of dissolved macronutrients (DIN and DIP) and level of primary productivity of a unit. Oligotrophic [Olig] <5 µg/L chlorophyll a; Mesotrophic [Meso] 5-50 µg/L chlorophyll a; Eutrophic [Eutr] >50 µg/L chlorophyll a. Alias: Water Column Parameters

Water_Mod: Energy Type: Wind [Wind]; Current [Curr]; Surface wave [SWav]; Internal Wave [IWav]; Tide [Tide]. Energy Intesity: Very low energy [VLEn], Low energy [LoEn], Moderate energy [MoEn], High energy [HiEn]. Energy Direction: Upward [Upwd]; Downward [Dnwd]; Horizontal [Hriz]; Baroclinic [Baro]; Seaward [Sewd]; Circular [Crcl]; Mixed [Mixd]. Primary Water Source: Watershed [Wshd], Local estuary exchange [LoEs], Local ocean exchange [LoOc], River [Rivr], Estuary [Estu], Marine [Mrin]. Anthropogenic Impact: Developed, Impounded/Diverted [Impd], Dredged area/Channel [Drdg], Filled [Flld], Contaminated [Cont]. Temporal Persistence: Stochastic [Stoc], Hours [Hors], Days [Days], Weeks [Weks], Months [Mnts], Seasons [Seas], Years [Yers], Inter-annual [IntA], Decades [Decs], Centuries [Cent]. Alias: Water Column Modifier

Water_Conf: The level of confidence in which a study was coded. Each column has been given a number indicating the confidence level. The number 1 is explicit and direct link, number 2 is inference from text, number 3 is extrapolated from other studies, number 4 is geographical associations, and number 5 is educated guess. Explicit and direct link: information within the study was explicitly stated. Inference from text: information within the study could be inferred but was not explicitly stated. Extrapolated from other studies: information was extrapolated from another study. Geographical association: information within the study is associated with its location. Educated guess: information based on the knowledge of the person coding the study. Alias: Water Column Confidence

CMECS_Code: All components of CMECS are intended to co-operate so as to describe the geological, chemical, physical, and biological associations that constitute habitat. The common code begins with System and Subsystem to describe a broad aquatic setting. The common code then defines strings from each of the five CMECS components. The code structure keeps all codes the same length and has a consistent format for the purpose of future programming or querying. Codes with an 'X' indicate missing data and with a '?' indicate data that cannot fit into the existing classification items of categories or subcategories. Certain systems have not yet assigned specified codes, for example, the water column component does not have a code for Fresh water. Following convention, assign Frsh for Fresh. In cases where system values have a range instead of a fixed value, use a dash to indicate the range, for example: Oligo-Meso. Alias: CMECS Code

CMECS_Stan: CMECS_Stan is a code for are standard attributes indicating physico-chemical, physical, spatial, geological, biological, anthropogenic, and biogeographic variables with defined categorical values that are used to classify or further describe CMECS types. The code structure keeps all codes the same length and has a consistent format for the purpose of future programming or querying. Codes with an 'X' indicate missing data and with a '?' indicate data that cannot fit into the existing classification items of categories or subcategories. In cases where system values have a range instead of a fixed value, use a dash to indicate the range. Alias: CMECS Standard Attribute Code

Notes: Any additional information regarding the study. Alias: Notes