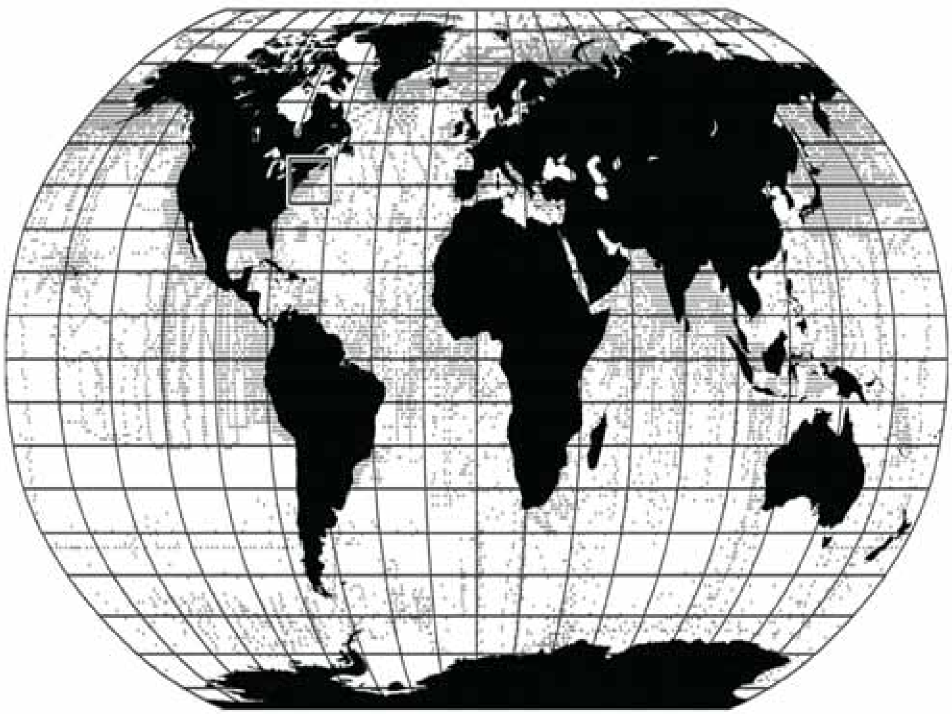
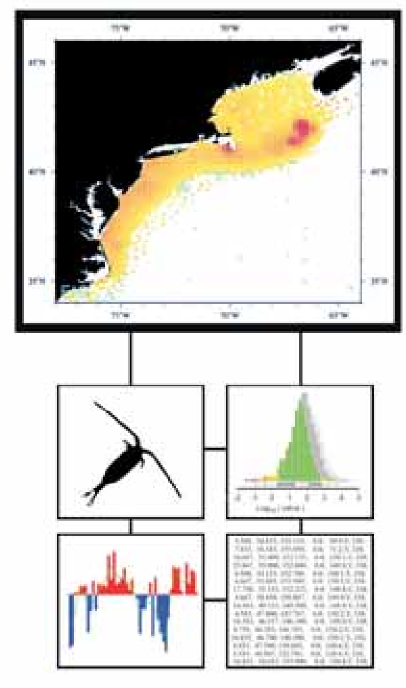
NMFS-COPEPOD

The **C**oastal & **O**ceanic **P**lankton **E**cology, **P**roduction & **O**bservation **D**atabase





COPEPOD: The Global Plankton Database

A review of the 2007 database contents and new quality control methodology.

2007

Office of Science & Technology
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

NOAA Technical Memorandum NMFS-F/ST-34
December 2007

About this document:

The mission of the National Oceanic and Atmospheric Administration (NOAA) is to understand and predict changes in the Earth's environment and to conserve and manage coastal and oceanic marine resources and habitats to help meet our Nation's economic, social, and environmental needs. As a branch of NOAA, the National Marine Fisheries Service (NMFS) conducts or sponsors research and monitoring programs to improve the scientific basis for conservation and management decisions. NMFS strives to make information about the purpose, methods, and results of its scientific studies widely available.

The NMFS Office of Science & Technology (NMFS-ST) uses the **NOAA Technical Memorandum** series to achieve timely dissemination of scientific and technical information that is of high quality but inappropriate for publication in the formal peer-reviewed literature. The contents are of broad scope, including technical workshop proceedings, large data compilations, status reports and reviews, lengthy scientific or statistical monographs, and more. NOAA Technical Memoranda published by the NMFS-ST, although informal, are subject to extensive review and editing and reflect sound professional work. Accordingly, they may be referenced in the formal scientific and technical literature.

This document should be cited as follows:

O'Brien, T.D. 2007. COPEPOD: The Global Plankton Database. A review of the 2007 database contents and new quality control methodology. U.S. Dep. Commerce, NOAA Tech. Memo., NMFS-F/ST-34, 28 p.

For further information direct inquiries to

Todd D. O'Brien
NOAA – NMFS – F/ST7
1315 East-West Highway - #12538
Silver Spring, Maryland 20910

Email: Todd.OBrien@noaa.gov

Phone: 301-713-2363 ext. 174
Fax: 301-713-1875

This document and the NMFS-COPEPOD database and data products are available online at:

http://www.st.nmfs.noaa.gov/plankton

COPEPOD: The Global Plankton Database

A review of the 2007 database contents and new quality control methodology.

Todd D. O'Brien

NOAA Technical Memorandum NMFS-F/ST-34
December 2007

U.S. Department of Commerce
Carlos M. Gutiérrez, Secretary

National Oceanic and Atmospheric Administration
Vice Admiral Conrad C. Lautenbacher, Jr., USN (Ret.)
Under Secretary for Oceans and Atmosphere

National Marine Fisheries Service
William T. Hogarth, Assistant Administrator for Fisheries

The *NMFS-COPEPOD* project was born on March 8, 2004, thanks to the efforts and foresight of Michael Sissenwine (retired, former *NMFS Director of Scientific Programs & Chief Science Advisor*). This project continues with the encouragement and support of Ned Cyr (*Chief, Marine Ecosystems Division, NMFS Office of Science & Technology*). Special thanks go out to Peter H. Wiebe (WHOI), Jack Jossi (NEFSC), Bill Peterson (NWFSC), and the members of the ICES Working Group on Zooplankton Ecology (http://WGZE.net) for their shared plankton expertise and support over the years.

You are not finished when you publish,
you are finished when you share the data.

The contents of this plankton database would not be possible without the effort and expertise of the plankton scientists and technicians that originally collected, identified, and quantified its plankton content. This database would also not be possible if these same scientists did not share the results of their hard work and expertise, namely the data itself. The goal of COPEPOD is to make these plankton scientists' hard work available to the rest of the scientific community in a standard, useable data format that also clearly identifies the original persons and entities that were responsible for collecting and preparing those data. Like a library, COPEPOD does not claim credit for its contents but strives to be valued for the breadth and comprehensiveness of the entire collection it offers to the public.

In exchange for using the contents of COPEPOD, we request three things of its users:

- Please acknowledge the original investigators of the data whenever possible;

  Guidance for how to cite the data is provided within each data set.

- Please acknowledge that the data were acquired online from;

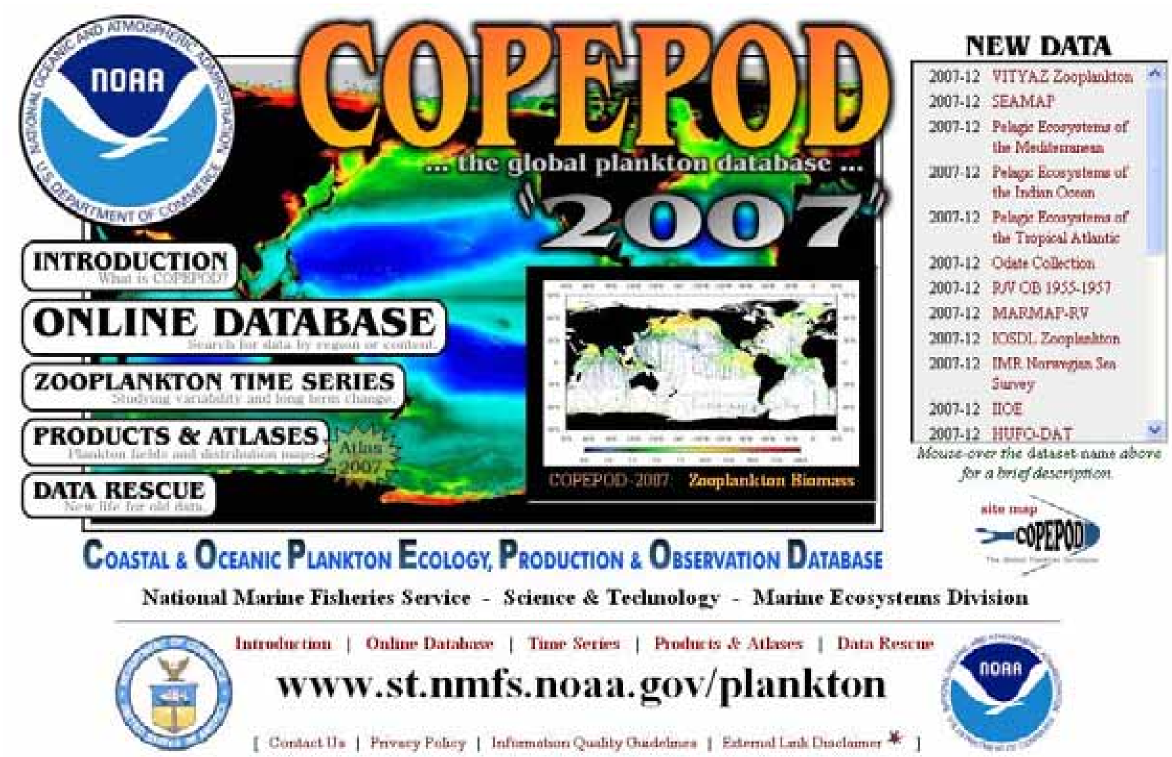
  COPEPOD: The Global Plankton Database (www.st.nmfs.noaa.gov/plankton).

- Please consider contributing your own data to this global compilation.

  More information is available online or by emailing Todd.OBrien@noaa.gov.

A large portion of the historical COPEPOD content was keyed from printed data reports and cruise summaries. COPEPOD is especially grateful for the help, support, and historical document and archive access granted by the staff of the NOAA Central Library (Silver Spring, Maryland), the staff of the WHOI-MBL libraries (Woods Hole, Massachusetts), and Charlotte Sazama of the World Data Center for Oceanography (Silver Spring, Maryland). COPEPOD's ongoing digitization of historical plankton data manuscripts is possible through funding from the NOAA Climate Data Modernization Program (CDMP).

Todd D. O'Brien
COPEPOD Project Leader
Todd.OBrien@noaa.gov



1. Introduction **1**

2. Data Sources **1**

2.1 The NMFS Ecosystem Surveys 2
2.2 Historical Plankton Data Search & Rescue 4
2.3 Data Centers, Institutions, & Project Data 4
2.4 Direct Investigator Submission 5

3. Data Processing **5**

3.1 Taxonomic Translation and Name Verification 5
3.2 Plankton Grouping 6
3.3 Calculation of Common Base-unit Values 9
3.4 Data Presentation & Access 10

4. "Quality Control" **12**

4.1 Value Categorization and Sub-Grouping 12
4.2 Ranging and Outliers 16
4.3 Data and Range Visualization 18

5. Data Products **21**

5.1 Online Database 21
5.2 Online Atlas 22
5.3 Global Zooplankton Biomass Fields 23

COPEPOD: The Global Plankton Database

Todd D. O'Brien
NOAA - National Marine Fisheries Service (NMFS)
Office of Science & Technology - Marine Ecosystems Division
Silver Spring, Maryland

The **C**oastal & **O**ceanic **P**lankton **E**cology, **P**roduction & **O**bservation **D**atabase (**COPEPOD**) is an online, global-coverage database of zooplankton abundance, phytoplankton abundance, and zooplankton biomass data. Based on over ten years of plankton data management and database development, the COPEPOD project focuses on providing a plankton-tailored data access interface, integrated plankton data products, and clear acknowledgment of the original plankton investigators. While new COPEPOD data are available online each month, this document summarizes the content of the database as of December 2007. This document also introduces an advanced technique for quality control and range checking of plankton data, now in use with all COPEPOD data content.

1. INTRODUCTION

In 2004, a new plankton database effort began, incorporating over ten years of plankton data management experience and user feedback into designing and building a new online data system designed specifically for plankton data and plankton scientists. The **C**oastal & **O**ceanic **P**lankton **E**cology, **P**roduction & **O**bservation **D**atabase (**COPEPOD**) now contains the entire *reprocessed* plankton content of O'Brien *et al.* (2002), the significant amounts of new plankton data presented in *COPEPOD-2005* (O'Brien 2005) and new data added since 2005. COPEPOD also represents a new approach to providing data access and investigator acknowledgement in a global-scale database. This new approach focuses on the individual data sets, highlighting each with a full summary of the exact content, sampling methods, and investigators associated with those data. By packaging these individual sets into data compilations and data products, a user can work at variety of local, regional, or global scales.

2. DATA SOURCES

COPEPOD-2007 content represents over ten years of plankton data compilation and processing. While some of the oldest content was available in early versions of the World Ocean Database (O'Brien *et al.* 2002, Conkright *et al.* 1998), this content was completely reprocessed before its inclusion in COPEPOD (*see O'Brien 2005*) and its continued presence in *COPEPOD-2007*. In both the 2005 and 2007 versions of COPEPOD, data from the ongoing and historical National Marine Fisheries Service (NMFS) sampling and the COPEPOD *Historical Plankton Data Search & Rescue* project (COPEPOD-SAR) dominate the database content (Figure 1).

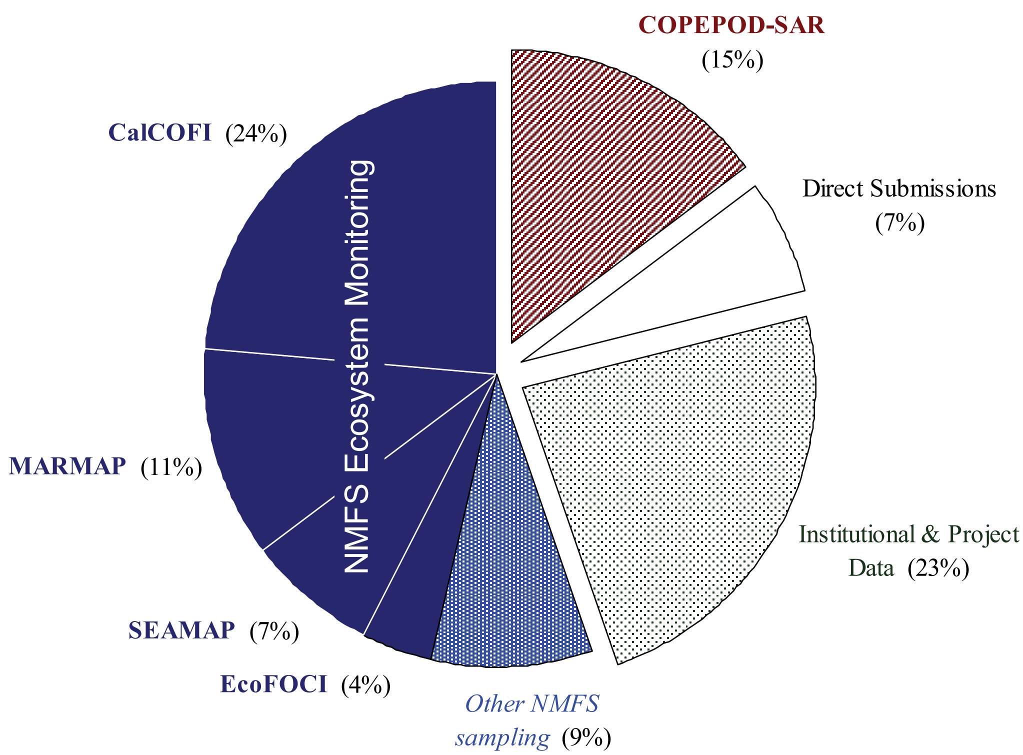


Figure 1: COPEPOD plankton data sources, shown as the percentage of all tows in the database, as of December 2007. NMFS associated sampling represent 55% of the database content.

The plankton content of COPEPOD comes from four main sources:

- Major ongoing NMFS Ecosystem Surveys (and other NMFS sampling programs);
- Historical Plankton Data Search & Rescue work (*COPEPOD-SAR*);
- Institutional & Project data;
- Direct Investigator Submission.

2.1 The NMFS Ecosystem Surveys

The National Marine Fisheries Service (NMFS) regularly samples plankton as part of its ongoing Ecosystem Survey programs (*Figure 2*). These regional sampling efforts include zooplankton displacement volumes as well as zooplankton and ichthyoplankton composition and abundance samples.

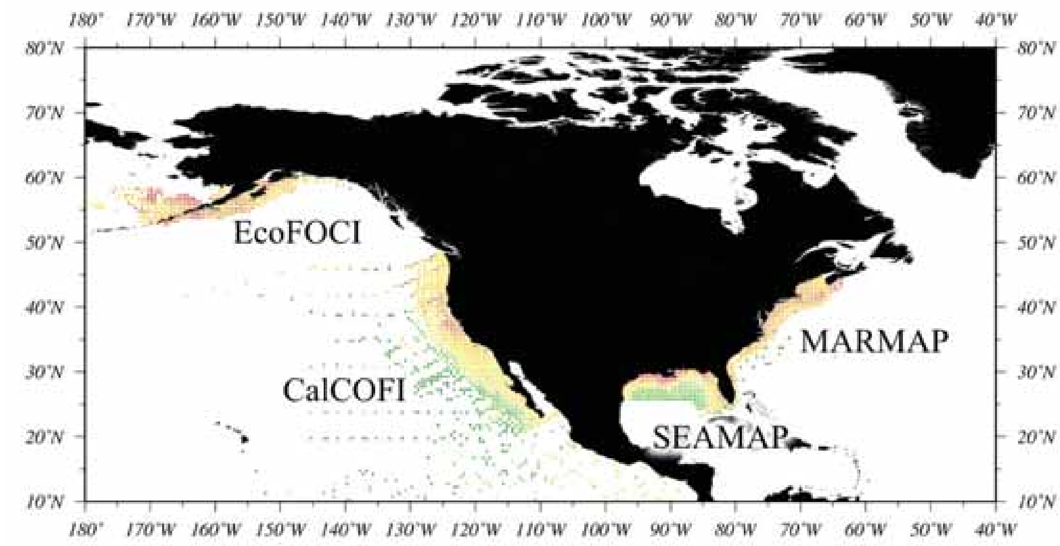


Figure 2: Map of zooplankton biomass and associated programs of the NMFS Ecosystem Surveys

CalCOFI - **Cal**ifornia **C**ooperative **O**ceanic **F**isheries **I**nvestigations (*1951- present*)
EcoFOCI - **Eco**systems & **F**isheries-**O**ceanography **C**oordinated **I**nvestigations (*1979-present*)
MARMAP - **Mar**ine **R**esources **M**onitoring, **A**ssessment, & **P**rediction (*1977-present*)
SEAMAP - **S**outh**Ea**st **M**onitoring & **A**ssessment **P**rogram (*1982-present*)

NMFS also has a rich history of zooplankton surveys and sampling going back to the early 1950's. Known in early years as the Bureau of Commercial Fisheries (BCF), NMFS/BCF worked together with State and research institutions to sample coastal US waters and Hawaii, and participated in large projects such as EASTROPAC which sampled far out into the equatorial Pacific (Figure 3).

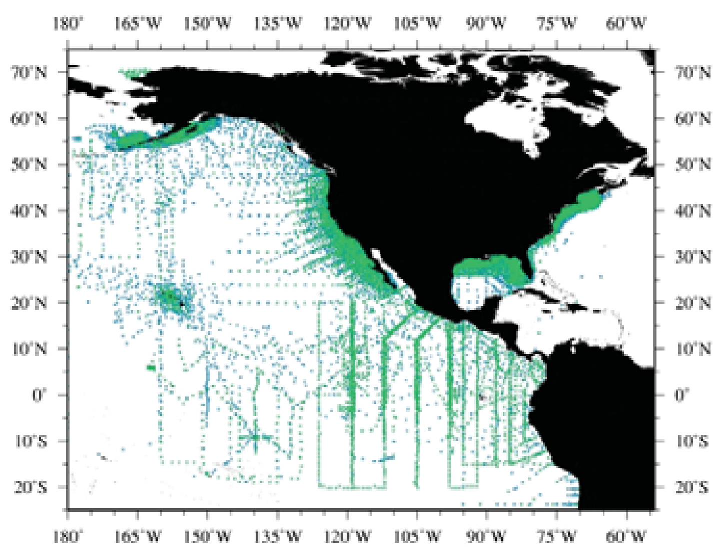


Figure 3: Map of historical and ongoing NMFS-associated plankton-sampling programs and surveys.

2.2 Historical Plankton Data Search & Rescue

Prior to 1998, large amounts of historical plankton data existed only in non-electronic, paper form. Some of these paper documents were available through journals or in library collections, but large amounts remained completely inaccessible, hidden in institutional archives or investigators' filing cabinets and storage rooms.

Over the last ten years, ongoing digitization efforts have managed to key large amounts of these data into an electronic form. In addition to the COPEPOD author's ten years of effort, other institutional and personal efforts have made large collections of historical Japanese (*e.g.,*Odate time series, Hokkaido University surveys) and Former Soviet Union (*e.g.,* Sergey Piontkovski's *Global Plankton Database* project) plankton data readily available to the scientific community.

The COPEPOD *Historical Plankton Data Search & Rescue* project (COPEPOD-SAR) is still actively keying these paper plankton data sources into a digital form, quality reviewing them, then distributing them online via the COPEPOD database. This past and ongoing keying work has been accomplished through funding and participation in NOAA's Climate Data Modernization Program (CDMP) and through previous funding from NOAA's Environmental Services Data and Information Management (ESDIM) program (Table 1).

Data Rescue Funding	Project Duration	Principal Investigator(s)
CDMP	2005 – present	*T.D. O'Brien*
ESDIM	2001 – 2004	*T.D. O'Brien*
ESDIM	1999 – 2000	*T.D. O'Brien & M.E. Conkright*

Table 1: Summary of COPEPOD plankton data search and rescue funding.

2.3 Data Centers, Institutions, & Project Data

Globally sampled plankton data were acquired in various electronic formats from international and regional data centers (*e.g.,* the British Oceanographic Data Centre, the Indian National Oceanographic Data Center, the Japanese Oceanographic Data Center, the U.S. National Oceanographic Data Center), research institutions (*e.g.,* Woods Hole Oceanographic Institute, Smithsonian Institution, Sir Alister Hardy Foundation for Ocean Science), and various oceanographic projects (*e.g.,* JGOFS, GLOBEC).

2.4 Direct Investigator Submission

An increasing amount of plankton data content is now being submitted directly to COPEPOD by the original collecting investigators. We especially welcome these data as the direct correspondence with the collecting investigator(s) allows us to insure we correctly represent the methods and data, and that we properly acknowledge the responsible investigators and institutions. As a result, these submissions tend to be the most recent and highest quality data in the COPEPOD database.

Whether it is a single cruise or a collection, COPEPOD invites more investigators to consider hosting their data through COPEPOD. Our goal is to provide quality data in an easy-to-access, easy-to-use database that also retains credit and acknowledgement for those whose hard work compiled the data.

Appendix I lists all *known* investigators whose work is present in *COPEPOD-2007* as of December 2007. Unfortunately, information on the collecting investigators or institutions is not always known or available. When possible we try to track down this information and add it to the database and content summaries. Likewise, COPEPOD welcomes feedback from the user community regarding missing or incomplete metadata.

3. DATA PROCESSING

The plankton data present in COPEPOD were originally distributed in hundreds of different formats (*e.g.,* tables, spreadsheets, ASCII text files). The first and major step in adding these data to a comprehensive database involves carefully translating these formats and variables into a common variable definition set and data structure. This step also involves reviewing the data documentation to ensure that all methods and metadata were accurately represented. Once this step is completed, the original data values are now available in an electronic database, but they do not necessarily have comparable (value) units or a common taxonomic indexing system which would allow them to be easily used by a database user. The second major step in building "COPEPOD" is to classify all of the plankton taxa into useable groups and to provide common base-unit values.

3.1 Taxonomic Translation and Name Verification

COPEPOD's policy on taxonomic name management is to preserve the original, investigator-provided, description as complete and correctly as possible. Spelling and taxonomic validity are checked against the Integrated Taxonomic Information System (ITIS, *http://www.itis.gov*), and stored in the COPEPOD database using both the ITIS-generated unique taxonomic serial number (TSN) and the ITIS-adopted spelling for that taxa. In cases of taxonomic synonyms, the original description is retained, but stored along with an additional TSN for the accepted/valid name. If a taxonomic name is not found in ITIS, it is assigned a temporary identifier (a sequence of unique, negative numbers, maintained by COPEPOD) until reviewed by and added to ITIS. A similar sequence of negative identifiers is also used to represent non-taxonomic or multi-taxonomic descriptions (*e.g.,* “phytoplankton”, “salps & doliods”, “jellyfish”). During the process, each taxonomic entry in COPEPOD is assigned a taxonomic quality flag (*e.g.,* “valid”, “synonym”, “awaiting ITIS review”, “invalid/unverifiable name”, “non-taxonomic group”).

Non-taxonomic identifiers and modifiers provided in the original description are also preserved in the COPEPOD database. These include life stage information (*e.g.,* “adult”, “juvenile”, “copepodite C3”), gender (*e.g., “male”, “female”*), size range (*e.g.,* “> 1 mm”, “2 – 5 um”), and ranking indicators (*e.g., “*spp.”, “other”, “miscellaneous”).

3.2 Plankton Grouping

The COPEPOD database contains over 1.5 million plankton abundance measurements, labeled with over 5,000 unique taxon identifiers that range from individual species names (*e.g., Calanus finmarchicus, Nitzschia delicatissima*) to general classes and families (*e.g.,* “copepods”, “euphausiids”). To allow for quick and simple access to all data from a general plankton grouping such as “diatoms” or “copepods” or even “phytoplankton”, a smart-index was added to each record within the COPEPOD database. (This smart-index is stored in addition to the original taxonomic descriptions and/or modifiers described in Section 3.1.)

The COPEPOD smart-index, called the Plankton Grouping Code (PGC), is a seven digit number which identifies the plankton taxa’s membership in up to four groups (Table 2). For example *Calanus finmarchicus*, a well known calanoid copepod, is assigned a PGC of “4212010”. This PGC indicates that *Calanus finmarchicus* is a “zooplankton” (Major Group = 4), a “crustacean” (Minor group = 21), a “copepod” (Focus Group = 20), and a “calanoid copepod” (Special Group = 10). Every and every database record for *Calanus finmarchicus,* along with any other “calanoid copepod” species in the database, will have this same “4212010” PGC code. This single PGC code can then be used to quickly identify and access a taxa group that contains hundreds of different species.

PGC	**Major Group (##○○○○○○)**	**Minor Group (○○##○○○○)**	**Focus Group (○○○○##○○)**	**Special Group (○○○○○○##)**	**Scientific Name**
1050000	Bacterioplankton	Cyanobacteria	-	-	*Oscillatoria thiebautii*
2160000	Phytoplankton	Diatom	-	-	*Skeletonema costatum*
42120**10**	Zooplankton	Crustaceans	Copepods	Calanoid copepods	*Calanus finmarchicus*
42120**10**	Zooplankton	Crustaceans	Copepods	Calanoid copepods	*Metridia pacifica*
42120**40**	Zooplankton	Crustaceans	Copepods	Cyclopoid copepods	*Oithona similes*
421**80**00	Zooplankton	Crustaceans	Euphausiids	-	*Euphausia pacifica*
4**32**0000	Zooplankton	Chaetognaths	-	-	*Parasagitta elegans*

Table 2: Examples of Plankton Grouping Codes (PGC) and inherent sub-grouping.

The four major, minor, and focus grouping levels of the PGC smart-index can be accessed by using simple *integer math*[*]. For example:

Major Groups can be selected by dividing any PGC by the 1000000:

If (PGC / 1000000) = 4 { it is a zooplankton }
If (PGC / 1000000) = 2 { it is a phytoplankton }

Example: "4212010" / 1000000 = 4 (*Calanus finmarchicus* is a zooplankton)

Minor groups can be selected by dividing any PGC by 10000:

If (PGC / 10000) = 216 { it is a diatom }
If (PGC / 10000) = 421 { it is a crustacean }
If (PGC / 10000) = 432 { it is a chaetognath }

Example: "4212010" / 10000 = 421 (*Calanus finmarchicus* is a crustacean)

Focus Groups can be selected by divided by 100:

If (PGC / 100) = 42120 { it is a copepod }
If (PGC / 100) = 42180 { it is an euphausiid }

Example: "4212010" / 100 = 42120 (*Calanus finmarchicus* is a copepod)

The active Plankton Grouping Codes used in *COPEPOD-2007* are listed in Table 3.

[*] ***A note on "integer math":***

In integer math, the result of any calculation is <u>rounded down</u> to the nearest integer. A real value result of "1.1", "1.7", or even "1.9999" will actually round down to "1". This is a fundamental property to remember when mathematically quantifying PGC group membership in software or programs. In most software and programming languages, enclosing the calculation in "INT()" will perform an integer-based calculation. For example:

INT(4286000/10000) = 428

PGC	MAJOR	Minor Groups	*Focus Groups*	*Special*
1000000	**Bacterioplankton**			
1050000		Cyanobacteria		
2000000	**Phytoplankton**			
2040000		Granuloreticulosa (foraminifera)		
2070000		Dinomastigota (dinoflagellates)		
2080000		Ciliophora (ciliates)		
2100000		Haptomonada (coccolithophores)		
2110000		Cryptomonada (cryptophytes)		
2120000		Discomitochondira (flagellates)		
2130000		Chrysomonada (chrysophytes)		
2160000		Diatoms		
2270000		Actinopoda (radiolarians)		
2280000		Chlorophyta (green algae)		
4000000	**Zooplankton**			
4030000		Cnidaria		
4032000			*Hydrozoa*	
4036000			*Scyphozoa*	
4038000			*Anthozoa*	
4040000		Ctenophora		
4050000		Platyhelminthes		
4090000		Nemertina		
4100000		Nematoda		
4130000		Rotifera		
4190000		Chelicerata		
4200000		Mandibulata		
4210000		Crustacea		
4211000			*Ostacoda*	
4212000			*Copepoda*	
4212010				*Calanoida*
4212030				*Harpacticoida*
4212040				*Cyclopoida*
4212050				*Monstriloida*
4212060				*Caligoida*
4213000			*Cirripedia*	

PGC	MAJOR	Minor Groups	*Focus Groups*	*Special*
4000000	**Zooplankton** (*continued*)			
4214000			*Mysidacea*	
4216000			*Isopoda*	
4217000			*Amphipoda*	
4217025				*Gammaridea*
4217050				*Hyperiidea*
4217075				*Caprellidea*
4218000			*Euphausiacea*	
4219000			*Decapoda*	
4220000		Annelida		
4225000			*Polychaetes*	
4230000		Sipuncula		
4260000		Mollusca		
4262500			*Gastropoda*	
4265000			*Bivalvia*	
4266000			*Scaphopoda*	
4267500			*Cephalopoda*	
4290000		Bryozoa		
4300000		Brachiopoda		
4310000		Phoronida		
4320000		Chaetognatha		
4330000		Hemichordata		
4340000		Echinodermata		
4342000			*Asteroidea*	
4344000			*Ophuiroidea*	
4346000			*Echinodea*	
4348000			*Holothuroidea*	
4350000		Urochordata		
4352500			*Ascidiacea*	
4355000			*Thaliacea*	
4355010				*Salps*
4355050				*Doliolids*
4357500			*Appendicularia*	
4360000		Cephalochordata		
5000000	**Ichthyoplankton**			

Table 3: Active *Plankton Grouping Codes* (PGC) present in *COPEPOD-2007*. Groupings are based on Margulis and Schwartz (1998).

.

3.3 Calculation of Common Base-unit Values

The original plankton measurements in COPEPOD came in a variety of different measurement units (*e.g.,* "number per ml", "number per 30-liter sample", "number per 100 m^3", "number per m^2", "number per total net sample"). Within COPEPOD, these values are stored as originally measured with the following minor adjustments:

- Bulk multipliers in the unit numerator or denominator were calculated out to be just "number per unit" (*e.g.,* "35 critters **per 1000 m^3**" = "0.035 critters **per m^3**"; "3.5 **x 10^4** critters **per m^2**" = 35,000 critters **per m^2**");

- For bacteria and phytoplankton data, base units (*e.g.,* "per ml", "per liter", "per m^3") were changed if necessary to keep the original values less than $1x10^9$ and greater than 0.00001 (*e.g.,* "bacteria = **1.5 x 10^9** per m^3" is stored as "bacteria = **1.5** per µL").

Even after these adjustments, the variety and types of original units still do not allow for easy inter-comparison of the data (*e.g.,* one can not immediately compare "#/m^3" to "#/total-net-sample"). To allow for easier use and inter-comparison of values, a *Common Base-unit Value* (CBV) was calculated and stored along with each original value. This CBV is available in both a "per-volume" (CBV-m3) and a "per area" (CBV-m2) format, assigned based on the plankton group and original measurement type (Table 4).

Measurement Type & Group	CBV-m3 unit	CBV-m2 unit
Biomass (wet mass, dry mass, AFDM)	mg / m^3	mg / m^2
Biovolume (displacement volume, settled volume)	ml / m^3	ml / m^2
Zooplankton Abundance	# / m^3	# / m^2
Phytoplankton Abundance	# / mL	# x 10^6 / m^2
Bacterioplankton Abundance	# / µL	# x 10^9 / m2
Ichthyoplankton Abundance	# / m^3	# / m^2

Table 4: Measurement types and units for Common Base-unit "per volume" (CBV-m3) and "per area" (CBV-m2) Values.

If the original value was already in the correct CBV-m3 or CBV-m2 units, no calculation was necessary to create that CBV type. Otherwise, the common base unit value was calculated using the metadata associated with that sample. For example, "per total-net-sample" (per haul) measurements used flow meter 'volume of water filtered' to calculate the CBV-m3. When volume of water filtered was not provided, the "volume of water filtered" was estimated by multiplying the mouth area of the net opening by the distance the net was towed through the water column. For vertical tows, this towing distance was the lower depth minus the upper depth. For horizontal tows, towing distance was estimated by using the average towing speed and tow duration. The method used to calculate the CBV-m3, CBV-m2, or volume filtered (if not provided) is stored in the database alongside each CBV value.

3.4 Data Presentation & Access

The COPEPOD approach to data management is to focus on individual data sets, highlighting each with a detailed summary of the exact data content, sampling methods, and investigators associated with those data. Each individual COPEPOD data set, called a “collection”, includes an inter-linked, multi-page, html-based graphics and text content summary (Figure 4, Figure 5).

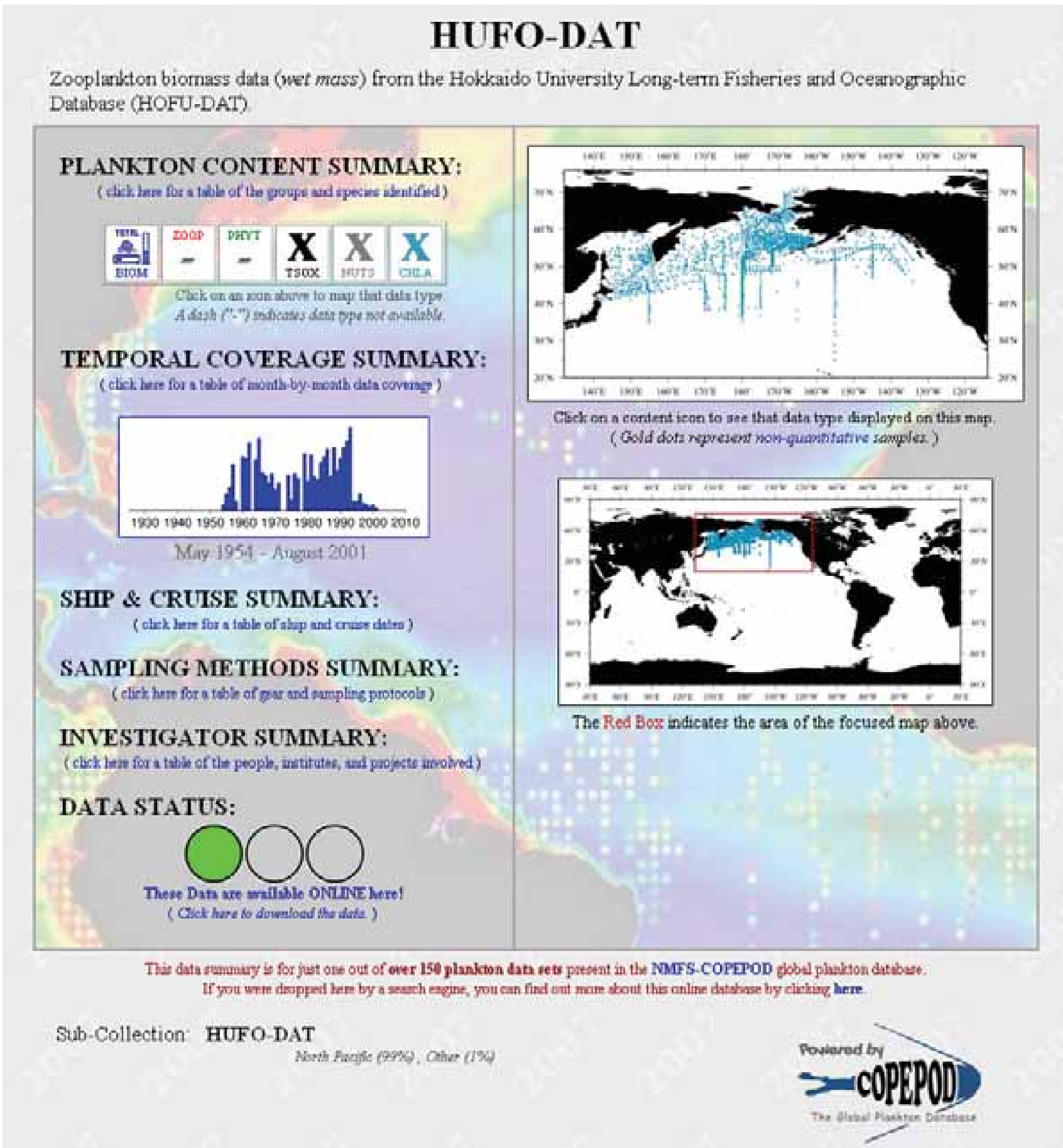


Figure 4: Example of a typical data collection summary main page for the Hokkaido University Long-term Fisheries & Oceanographic Database (HUFO-DAT).

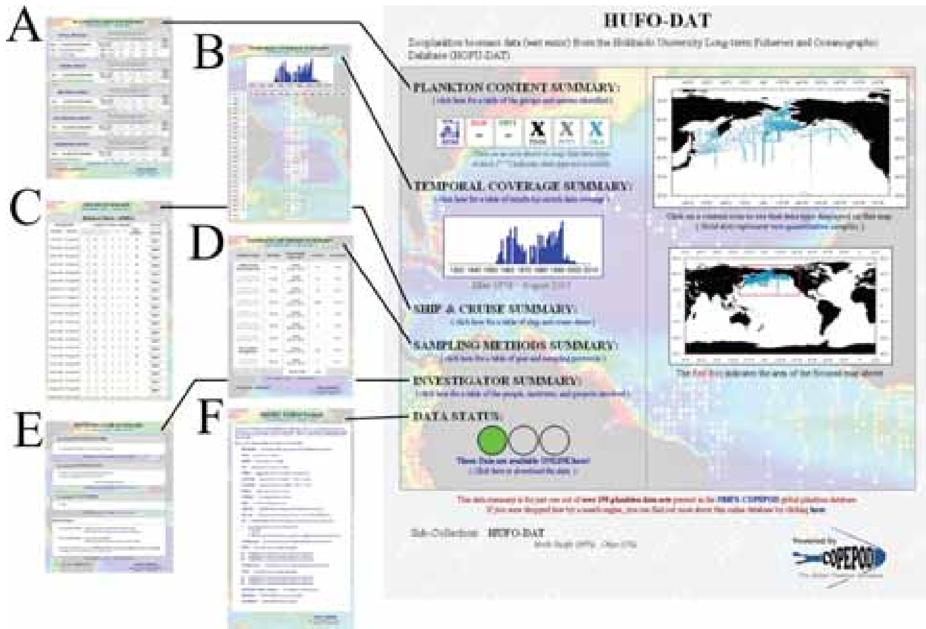


Figure 5: Examples of linked, sub-page summaries available within each COPEPOD data collection.

Each Collection Summary contains the following information and sub-pages:

A The **PLANKTON CONTENT SUMMARY** features a data distribution map on the main page and links to tables listing all of the plankton groups present in the collection. Additional links in this table provide lists of all species identified in the data collection.

B The **TEMPORAL COVERAGE SUMMARY** features a data-by-years histogram on the main page and links to a table listing the number and month-by-month composition of the data collection for each year.

C The **SHIP & CRUISE SUMMARY** links to a table of all ships, cruise dates, and variables measured during each cruise present in the data collection.

D The **SAMPLING METHODS SUMMARY** links to a summary of sampling and processing methods along with a table that summarizes the gear and mesh sizes used in the data collection.

E The **INVESTIGATOR SUMMARY** links to a table that lists the known investigators, institutions, countries, projects, and data sources associated with the data collection.

F The **DATA STATUS** summary features a "stop light" icon that indicates the availability of the data: A "green light" indicates that the data are currently available online; a "yellow light" indicates the data are in final review but available for early release by request; and a "red light" indicates that the data are available by contacting the original investigator. This section also features a link that provides download and data format options.

4. "Quality Control"

One of the biggest challenges to building a database of plankton sampled with a wide variety of sampling methods and sampling gear is checking the quality of the data. One approach is to do an intensive review of the entire database every few years and then release it after that review is completed (*e.g.,* World Ocean Database 1994, 1998, 2001, 2005). While this approach is thorough, it is time intensive, must be repeated "from scratch" for each new review, and it greatly slows the frequency of new data release. With its monthly data release schedule, this approach was not a possibility for COPEPOD. What was needed was a (value) range checking system that could quickly compare new plankton data to the thousands of other plankton data already present in the main database. After that automated review, the new data could be immediately released and the new data values could be added to the range checking "data pool" to improve future ranging checks. Using this technique, the ranging data pool is constantly improving as new data are added to the collection, so one would benefit from periodically re-checking the older data sets to incorporate the improved ranging. This automated system is now used in COPEPOD, with advanced ranging flags present in all *COPEPOD-2007* data collections.

In terms of plankton data, the main purpose of "quality control" is to check for errors in the database incorporation process versus quality control of the original data. In general, plankton data are usually "correct" in their original source media and any anomalous values found in these data are due to natural processes (*e.g.*, blooms, swarms, patchiness) or mechanical sampling issues (*e.g.*, gear failure or clogged nets). The original authors often annotated these mechanical or bloom events within the original data documentation or data tables, but these annotations may not have been passed along when the data were later digitized and/or added to other databases. The process of putting these data into the database itself is typically the biggest reason for errors in the data, ranging from metadata mis-translation (*e.g.,* in the foreign document, did the author mean "*millimeters*" or "*micrometers*" with the label "mm"), mislabeling of data types (*e.g.,* the data tables say "per m^3", but the documentation says "per 1000 m^3"), and a variety of numeric uncertainties (*e.g., "*Is the comma in "1,234" a thousands indicator or a decimal indicator?"). In each of these examples, the value ranging question is not "Is this value 5.6 or 5.7?" but rather "Is this value 5.6 or 5600?". These large differences are fairly easy to detect with automated ranging checks if the system is correctly comparing equivalent data types (*e.g.,* "comparing apples to apples, and oranges to oranges").

4.1 Value Categorization and Sub-Grouping

The distribution and concentration of plankton in the water column varies by region, by season, and even by time of day. The sampling gear and methods used to collect the plankton also play a major role in exactly how much and what members of the plankton community are actually captured. Finally, how these samples are processed varies from investigator to investigator, with some investigators processing the sample to each individual species (*e.g.,* "100 *Calanus finmarchicus*", "50 *Acartia longiremis*", "50 *Calanus* other") and others only processing to broad groups (*e.g.,* "200 copepods"). Any attempt to set general ranges for these data values must therefore begin by taking into account factors such as the net mesh size, the season or month, the oceanographic region, and the taxonomic resolution and binning of the samples.

COPEPOD currently contains over 2 million plankton biomass and abundance values, with observations from over 5,000 taxa, sampled by hundreds of different gear and mesh sizes in various regions around the world. While *COPEPOD-2005* and *WOD-2001/2005* used very basic value classes for simple range checking, *COPEPOD-2007* introduces the next generation of plankton data range checking through advanced categorization and grouping (Table 5).

	WOD-2001 / 2005* (O'Brien *et al. 2002)*	**COPEPOD-2005** (O'Brien 2005)	**COPEPOD-2007**
Taxonomic Resolution	3 broad taxa groups *("bacteria / phyto / zoo")* + 5 biomass types	21 Major BGC Groups + 5 biomass types	85 MajorMinor PGC groups + 7 taxa-hierarchy levels + 4 life stage categories +6 biomass types
Spatial Resolution	Global	Global	Global + 15 regions
Temporal Resolution	None *(all data, "annual")*	"Annual" + 4 seasons	"Annual" + 4 seasons + Day/Night
Mesh Sizes	None (all data)	None (all data)	6 mesh size categories

* WOD-2005 continued use of the WOD-2001 quality control system.

Table 5: History and complexity of plankton data quality control methods in WOD and COPEPOD.

COPEPOD-2007 Sub-Categories:

The first step in the *COPEPOD-2007* advanced plankton quality control and value ranging system is to separate each and every plankton value into a series of sub-categories that best represents the actual value type and sampling methods used for that value. The following sub-categories were used for ranging the *COPEPOD-2007* data:

Plankton Group: As mentioned in Section 3.2, each plankton observation is assigned to its corresponding plankton groups, such as "diatoms" or "copepods" or "chaetognaths", and assigned a Plankton Grouping Code (PGC) identifier. The *COPEPOD-2007* ranging system currently divides the plankton taxa into 85 different plankton groups.

Life Stage: Each taxonomic observation is split into one of four life stage categories. These categories are "adult", "larger sub-stages", "smaller sub-stages", and "eggs". The main purpose for these groupings is to separate out the smallest life stage counts. While these smaller life stages may be counted in larger mesh nets, they generally require smaller mesh sizes and additional microscope time and expertise to properly enumerate.

Taxonomic Hiearchy Level: This is a very important sub-category as counts from a single species will clearly have a different value range than a count (sum) of an entire class or order (composed of multiple species). Each taxa observation is split into seven taxonomic hierarchy categories: 1= species, 2= genus, 3= family, 4= order, 5 = class, 6 phylum, and 7=kingdom. "Super-", "infra-", and "sub-" extensions are ignored (*e.g.,* "suborder" = "super-order" = "order" = 4). Regardless of the hierarchy category, any taxa name with a "sp." species indicator (*e.g., Calanus* sp., Copepod sp., Crustacean sp., Diatom sp.) is automatically assigned to level "1" (species level) as it is assumed to refer to a single, unidentified species from that group.

Geographic Region: Each data value is checked against other data from the same geographic region as well as all data from the entire world ("global"). The geographic regions used in *COPEPOD-2007* (Figure 6) are based on those used by the *World Ocean Database* (*e.g.,* Conkright *et al. 2002,* O'Brien *et al.* 1998), without coastal/open-ocean designations.

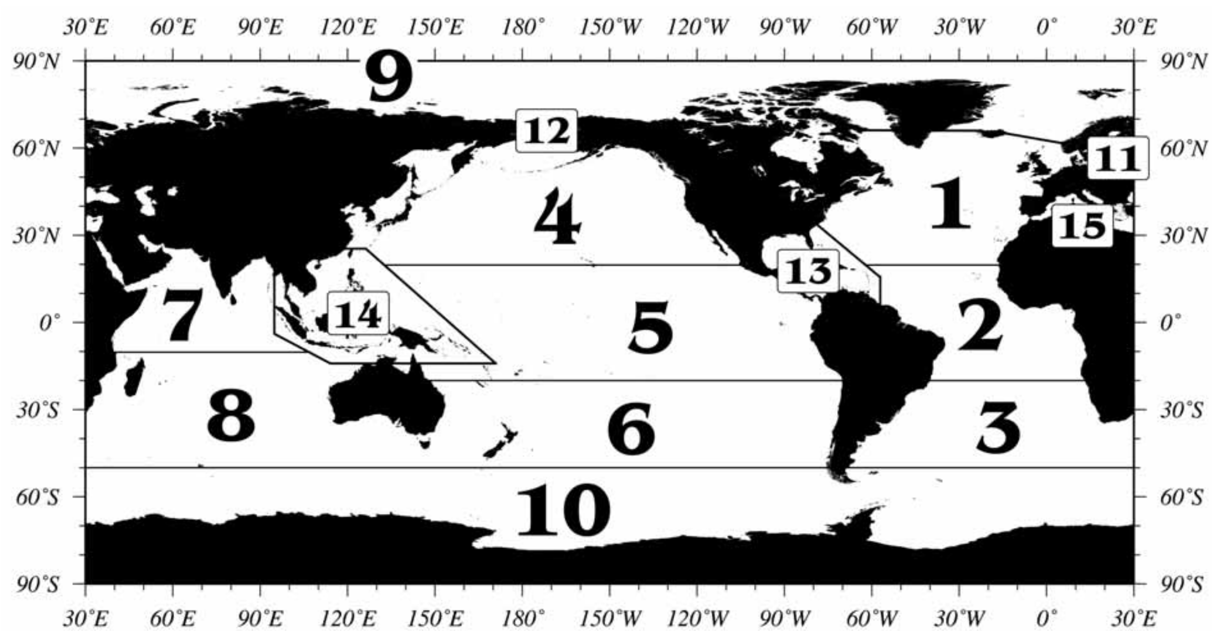


Figure 6: Map of geographic regions used by the *COPEPOD-2007* ranging system.

(1) North Atlantic, (2) Equatorial Atlantic, (3) South Atlantic, (4) North Pacific, (5) Equatorial Pacific, (6) South Pacific, (7) North Indian, (8) South Indian, (9) Arctic, (10) Antarctic, (11) Baltic, (12) Bering Sea, (13) Gulf of Mexico & Caribbean, (14) Indonesia, (15) Mediterranean.

Seasons: Each data value is checked against other data from the same "season" as well as all data from "any season or month" (annual). *COPEPOD-2007* seasons are defined as "winter" (December - February), "spring" (March - May), "summer" (June - August), and "autumn" (September - November).

Time of Day: Each value is checked against other data from the same "day" or "night" category, where "night" is defined as one-hour after local sunset* until one-hour before local sunrise* and "day" is defined as one-hour after sunrise until one-hour before sunset. The one hour buffers surrounding sunrise and sunset were excluded to avoid samples taken during the day/night transition. Data from these transition periods, as well as any samples or data sets with missing times, were checked against a general "any time / all times" sub-category (instead of "day" or "night").

* Local sunrise and sunset times are calculated using sample latitude, longitude, date, and time in the United States Naval Observatory Sunrise/Sunset Algorithm (*Almanac for Computers*, 1990).

Net Mesh Size: Each observation is checked against other data from the same general net mesh size. *COPEPOD-2007* mesh size categories were based upon the frequency of employed mesh sizes in the main database and labeled per general target organisms. Raw mesh sizes were split into one of six size categories: Meso-zooplankton mesh sizes (200 µm, 300 µm, 500 µm), fine plankton mesh sizes (050 µm, 100 µm), and "bottle". Table 6 illustrates the mesh size categories and criteria used within *COPEPOD-2007*.

		Fine Mesh		Meso Mesh		
	"bottle"	"050 µm *25 - 76 µm*	"100 µm *94 - 125 µm*	"200 µm *150 - 253 µm*	"300 µm *270 - 417 µm*	"500 µm *470 - 570 µm*
Total Biomass	-	X	X	X	X	X
Bacterioplankton	X	*	*	*	*	*
Phytoplankton	X	X	X	*	*	*
Zooplankton	-	X	X	X	X	X
Ichthyoplankton	-	-	X	X	X	X

Table 6. Range checking mesh categories used within *COPEPOD-2007*.

* Phytoplankton from mesh sizes > 100 µm and non-bottle bacterioplankton are flagged as "gear bias" (*i.e.,* specimens periodically snagged on the net but generally smaller than the net mesh opening).

Combining Sub-Groups

In addition to splitting the data into sub-categories, the *COPEPOD-2007* ranging system combines individuals into "combined grouping sums" (CGS), to allow for additional ranging checks at higher group or taxa-hierarchy levels. For example: A new data collection has comprehensive individual copepod genus and species-level counts, but no other genus or species data are available for this specific region or season in the main database. The main database does, however, have thousands of "total copepod" counts in that same region and season. To see if the new data have reasonable value ranges, the COPEPOD ranging system can add together all PGC "copepod" counts within the new data to create a combined group sum (CGS) for "total copepods". This CGS value can then be compared to the total copepod data, already present in the main database, to check for any gross value errors in the new data.

Future Sub-Categories

The COPEPOD ranging system is still in a developmental state. What is presented above is actively being applied to all *COPEPOD-2007* data, but additional sub-categories are being added and tested for incorporation in the near future. These improvements include:

MONTHLY RANGES: While the addition of "seasons" allows for tighter temporal ranging, the definition of a season actually varies with latitude and geographic region. The addition of month-by-month ranging will greatly improve the temporal resolution of the ranging system.

SAMPLING DEPTH: The majority of the plankton data in COPEPOD are from vertical or oblique tows sampling from around 200 meters depth (or the bottom, if it is shallower) to the surface. Additional depth categories will be added to handle surface-only tows and/or depths to 500 meters or deeper.

IMPROVED GEOGRAPHIC REGIONS: In most regions, there is a visable difference between the productive near-shore ("coastal") area and the typically oligotrophic open-ocean areas (*see* Figure 11 *in the Data Products section for an example*). In the same figure, we also see that regions such as the Equatorial Pacific have large differences between the western and eastern side of the basin. Incorporating "near shore vs. open ocean" sub-regions, or switching to a grid of standard geographic cells (*e.g.,* 5° x 5° latitude-longitude boxes*),* would better separate these productive and oligotrophic regions, allowing for tighter value ranging checks.

4.2 Ranging and Outliers

For each ranging sub-category, simple statistics are run on all available COPEPOD data within that sub-category and used to create a "ranging set" for that sub-category. For example, ranging statistics for "North Pacific total wet mass (mg/ m^3) data sampled with a 300 µm mesh net" might look like Table 7.

PGC	Region	Season	Mesh	"n"	< 99.99%	< 99.9%	< 99%	Median	> 99%	> 99.9%	> 99.99%
-403	4	13	300	**2,894**	0.01	0.04	0.60	16.8	242	707	1,000
-403	4	14	300	**5,305**	0.01	0.25	0.84	58.7	545	979	1,200
-403	4	15	300	**12,202**	0.04	0.42	1.68	80.5	1,752	6,253	13,111
-403	4	16	300	**5,469**	0.01	0.13	0.42	37.7	298	730	1,769

Table 7: Example of ranging sets for Total Zooplankton Wet Mass data (*PGC = -403*) sampled with 300 µm mesh in the North Pacific (*Region = 4*). Shown are ranging sets for general winter (*Season = 13*), spring (*14*), summer (*15*), and autumn (*16*) time periods.

Each ranging set contains the minimum and maximum value limits for a data value to be considered within 99%, 99.9%, or 99.99% of all the other available data in the main database. These ranging limits were selected so each category was ten-times larger than the previous category. If a data value falls outside of one of these ranging limits, that data value is flagged. The actual ranging flag assigned is dependent on a minimum "n" (the number of other data values available in the main database for that sub-category). At least 100 values are needed in a sub-category for any ranging flag to be assigned. If 1,000 or more values are available, the tighter 99.9% ranging check can be assigned, and the even tighter 99.99% ranging flag can be assigned if 10,000 or more values are available. This "n"-based ranging and flagging is summarized in Table 8.

	2007 Flag	"n" range *weaker* **stronger**
"Gear Bias"	-9	Applies to bacteria and phytoplankton (see Table MM).
< 99.99% of all data	-7	"n" > 100,000
	-6	*"n" > 10,000*
< 99.9% of all data	-5	"n" > 10,000
	-4	*"n" > 1,000*
< 99% of all data	-3	"n" > 1,000
	-2	*"n" > 100*
"Zero Value" indicator	-1	*Used for "zero values".*
"reasonable value"	**0**	
"n < 100" indicator	1	"n" < 100 (no ranging)
> 99% of all data	2	*"n" > 100*
	3	"n" > 1,000
> 99.9% of all data	4	*"n" > 1,000*
	5	"n" > 10,000
> 99.99% of all data	6	*"n" > 10,000*
	7	"n" > 100,000

Table 8: Table of *COPEPOD-2007* ranging flags and criteria.

The *COPEPOD-2007* ranging system assigns three ranging flags to each value. The first flag is a "global – annual" comparison flag, in which same sub-category data from all ocean regions and any season are used for the check. The second flag is a "regional – annual" comparison flag, in which same sub-category data from the same region (but any season) are used for the check. The third flag is a "regional – seasonal" comparison flag, in which same sub-category data from the same region and same season are used for the check. When monthly checks are added, a fourth flag will be added for "regional – monthly" comparisons.

General Comments about flag results:

- The purpose of the ranging flags is to allow for a quick and general indication of how a data value compares to similar data values in the entire database. The data are flagged, but not removed, allowing the data user to use or ignore the flags by their own choice.

- The higher the flag value (*or the lower, in the case of negative flags*), the more likely it is that the data are non-representative (anomalous). For example, a flag value of "7" means that this specific data value is larger than 99.99% of at least 100,000 other existing data values for the exact same plankton measurement. While it could still be a legitimate value, it is probably worth investigating before using it.

- It is possible for value to have different ranging flags assigned for each of the three flags. For example, a higher spring value may be flagged as a ">99%" global value, but be perfectly reasonable within its regional and seasonal sub-categories. Likewise, a specific basin or season may have limited data, so a weaker or "n <100" flag may be assigned for the regional or seasonal check, but a stronger flag could be assigned for the global check.

- Having a few values outside of the 99% limit is usually okay, especially if the bulk of the other data falls within those 99% limits. If large portions of the data flagged as >99%, or if multiple >99.9% or>99.99% flags are present, the data would be investigated.

4.3 Data and Range Visualization

The COPEPOD ranging system generates standard results tables for each data sub-category. Each table reports all plankton groups present in the data collection, listing the number of observations and percentage of those observations flagged for each flag type, similar to Table 9.

PGC	Region	Season	Mesh	*"n"*	**<99.99%**	<99.9%	*<99.0%*	**"okay"**	*>99.0%*	>99.9%	**>99.99%**
-403	**WORLD**	***Annual***	300	485	.	0.60%	.	99.00%	*0.40%*	.	.
-403	*N.Pacific*	***Annual***	300	485	.	0.60%	.	99.00%	*0.40%*	.	.
-403	*N.Pacific*	*Winter*	300	47	.	.	.	**100.00%**	.	.	.
-403	*N.Pacific*	*Spring*	300	90	.	3.30%	.	95.60%	*1.10%*	.	.
-403	*N.Pacific*	*Summer*	300	309	.	.	*0.60%*	99.00%	*0.30%*	.	.
-403	*N.Pacific*	*Autumn*	300	39	.	.	.	97.40%	*2.60%*	.	.
4320000	**WORLD**	***Annual***	300	117	.	.	.	**100.00%**	.	.	.
4320000	*N.Pacific*	***Annual***	300	117	.	.	.	**100.00%**	.	.	.
4320000	*N.Pacific*	*Winter*	300	0	.	.	.	.	.	.	.
4320000	*N.Pacific*	*Spring*	300	0	.	.	.	.	.	.	.
4320000	*N.Pacific*	*Summer*	300	117	.	.	.	**100.00%**	.	.	.
4320000	*N.Pacific*	*Autumn*	300	0	.	.	.	.	.	.	.

Table 9: Example of COPEPOD ranging results for *Total Wet Mass* biomass data (*PGC = -403*) and *Chaetognath* abundance data (*PGC= 4320000*) sampled in the North Pacific with ~300 µm mesh nets.

The results tables can be annotated for quick, “at-a-glance” reviewing. In Table 9 the flagged-data columns are color-coded with “okay” values shown in green and flagged values are shown in gray or red. One can quickly scan through the entire results file, looking for instances of red to investigate. In this example, the Chaetognath data are all “okay” (green), all values fall within the 99% range, but some of the Total Wet Mass data have really low values (red, “<99.9%” values). There were 485 total wet mass tows in this example, most of them (309) made in the summer. While some data were flagged as >99% or <99% (gray), they are only a small portion of then entire data set and seem reasonable. The red flags of “< 99.9%”, however, should be examined (*see below for further discussion*).

While Table 9 is useful for quickly reviewing a data set, plotting and viewing the same results can improve the ability to quickly review the data. Figure 7 is an example of a visualization for the results shown in Table 9, with additional information conveyed in the sub-tables.

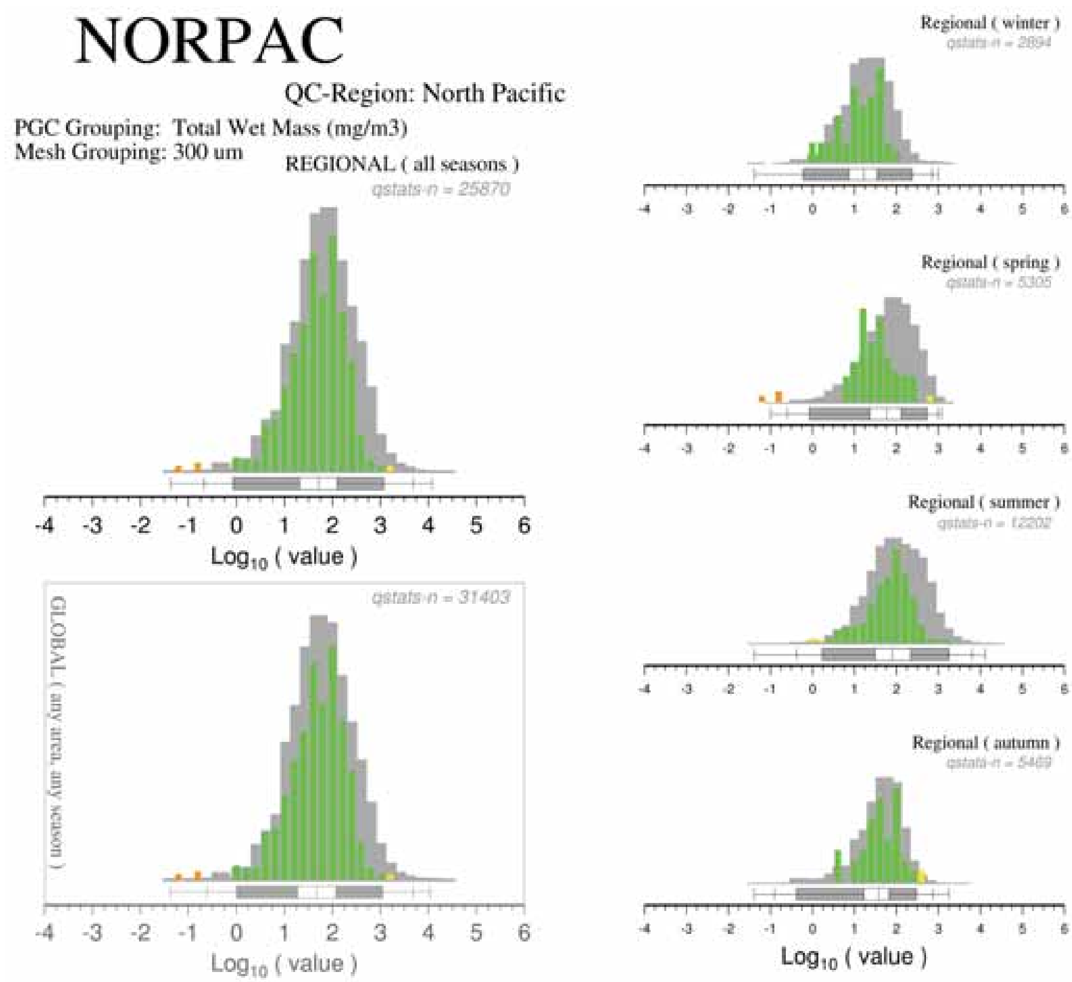


Figure 7: Ranging visualization for North Pacific *Total Wet Mass* data from the NORPAC project. The green/yellow/orange bars show NORPAC value distributions. The gray bars show value distributions for all corresponding wet mass data present in the entire COPEPOD database (“*qstats-n*” indicates the number of these values).

An explanation of Figure 7 (shown on previous page):

Bottom left sub-figure: Comparison of the NORPAC data to over 30,000 other "300 µm mesh wet mass values" (from any region, any season) present in the COPEPOD database.

Top left sub-figure: Comparison of the NORPAC data to over 25,000 other "300 µm mesh North Pacific wet mass values" (from any season) present in the COPEPOD database.

***Right column of four sub-figures*:** Comparison of the NORPAC data to other North Pacific ~300 µm mesh wet mass values from the same season.

In each figure, the box-n-whiskers portion immediately below the histograms indicates the ranging limits for 99% (gray outer box), with whiskers indicating the 99.9% (inner tick) and 99.99% (outer tick) limits. The white central box indicates the range of 50% of all data, the inner-most tick indicating the median of all data values.

The histograms in Figure 7 show the frequency of $\log_{10}$ total wet mass values. The two orange bars (seen in the global-annual, regional-annual, and regional-spring plots) are roughly 1000 times ($\log_{10}$ = "3") smaller than the bulk of the other data values. (These are the same values indicated with the red "<99.9%" in Table 9.) Upon checking the original data source, one would find an asterisk (*) next to the low values indicated gear problems for the three tows making up these points.

In general, the isolated values caused by a tow through a bloom, or equipment failure, will show up as individual spikes (as seen in the orange bars of Figure 7). Systematic errors, where an entire group of values is too high or too low due to mistranslated units or decimal errors, will typically show up as a shift in some or all of the values in the histogram. For example, the red values in Figure 8 turn out to be from the same cruise of a multi-cruise project. In the documentation for that cruise was a note that all data from that cruise were "per 100 m^3", a change from all previous cruises in the same project. By missing this one-time adjustment to the units, the data were accidentally loaded at 100 times their correct value. When plotted on the $\log_{10}$ scale, this 100x error appears as a $\log_{10}$ shift of "2" as seen in the red cluster in Figure 8.

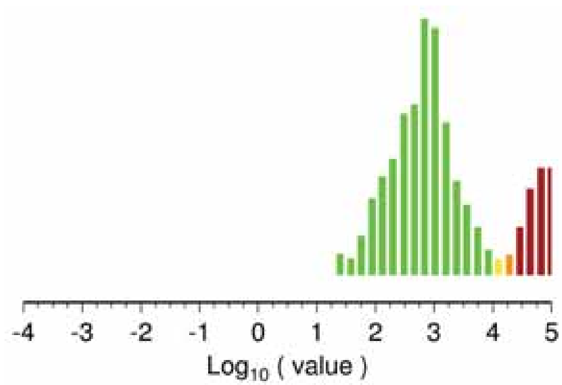


Figure 8: Example of a visual shift caused by systematic errors during data translation.

5. DATA PRODUCTS

In addition to offering hundreds of individual data collections, COPEPOD also offers regional and taxonomic compilations and data products which allow a user to work with the data at a variety of geospatial (*e.g.*, local, regional, or global) and data processing levels (*e.g.,* raw individual, species compilations, gridded aggregates and analyzed mean values).

5.1 Online Database

The online COPEPOD database (http://www.st.nmfs.noaa.gov/plankton) is the entry point for access to all of the COPEPOD data collections, compilations, and data products. Data collections can be easily searched by their associated geographic regions, projects, research vessels or cruise, or even credited investigators. Data are available in both an abbreviated "spreadsheet-friendly" and a comprehensive "all-information" format.

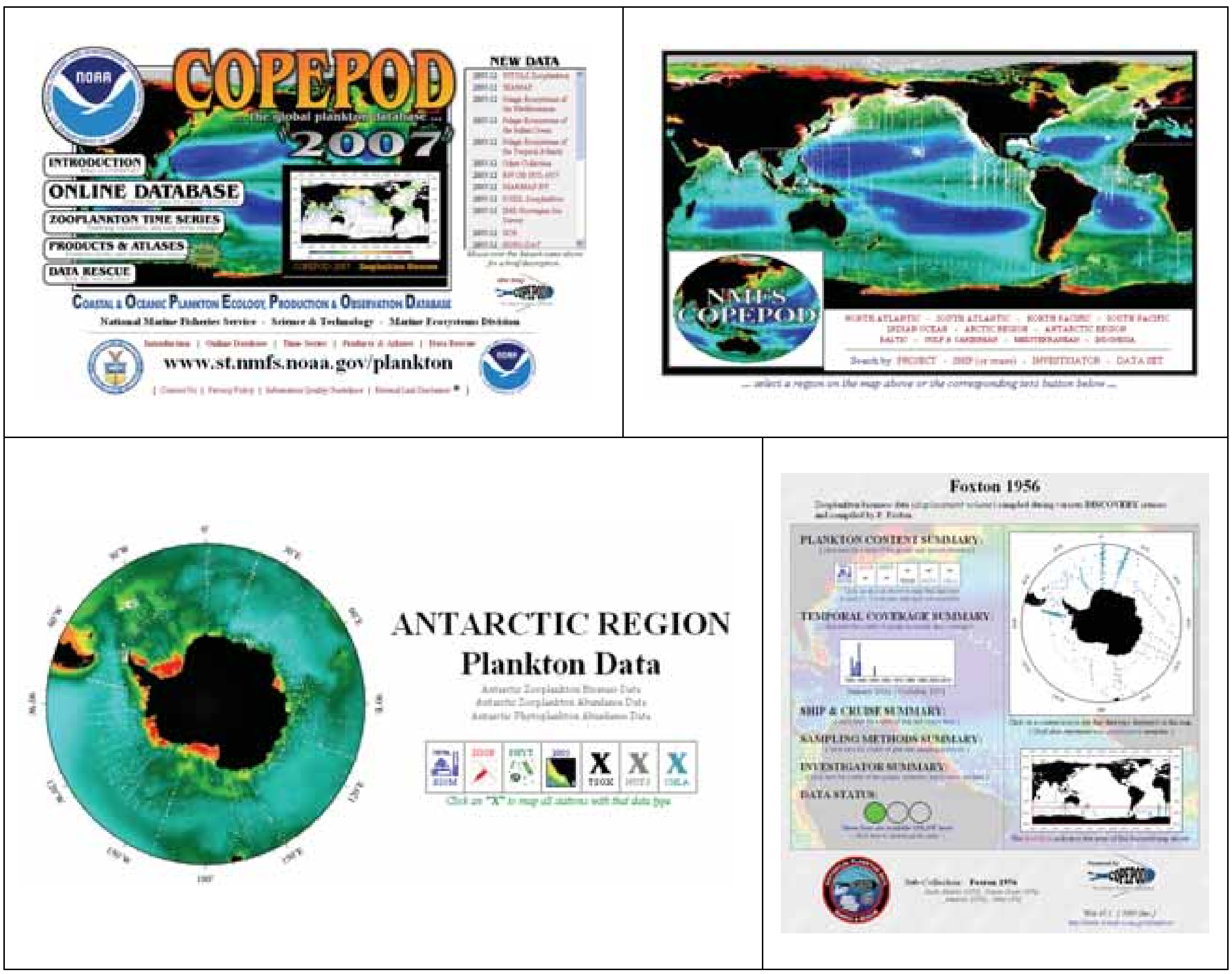


Figure 9: The COPEPOD online database features a variety of searching and access options.

5.2 Online Atlas

COPEPOD-2007 features an electronic atlas of data distribution maps and species lists for each of the major Plankton Groups shown in Table 3. Within each map, a graphical dot indicates the presence of at least one tow (or bottle sample) in the main database which has observations for that plankton group. A summary table lists the name of all species or measurement types present in the database, listed in order of frequency. Finally, users can download a global compilation containing all available data for that specific plankton group.

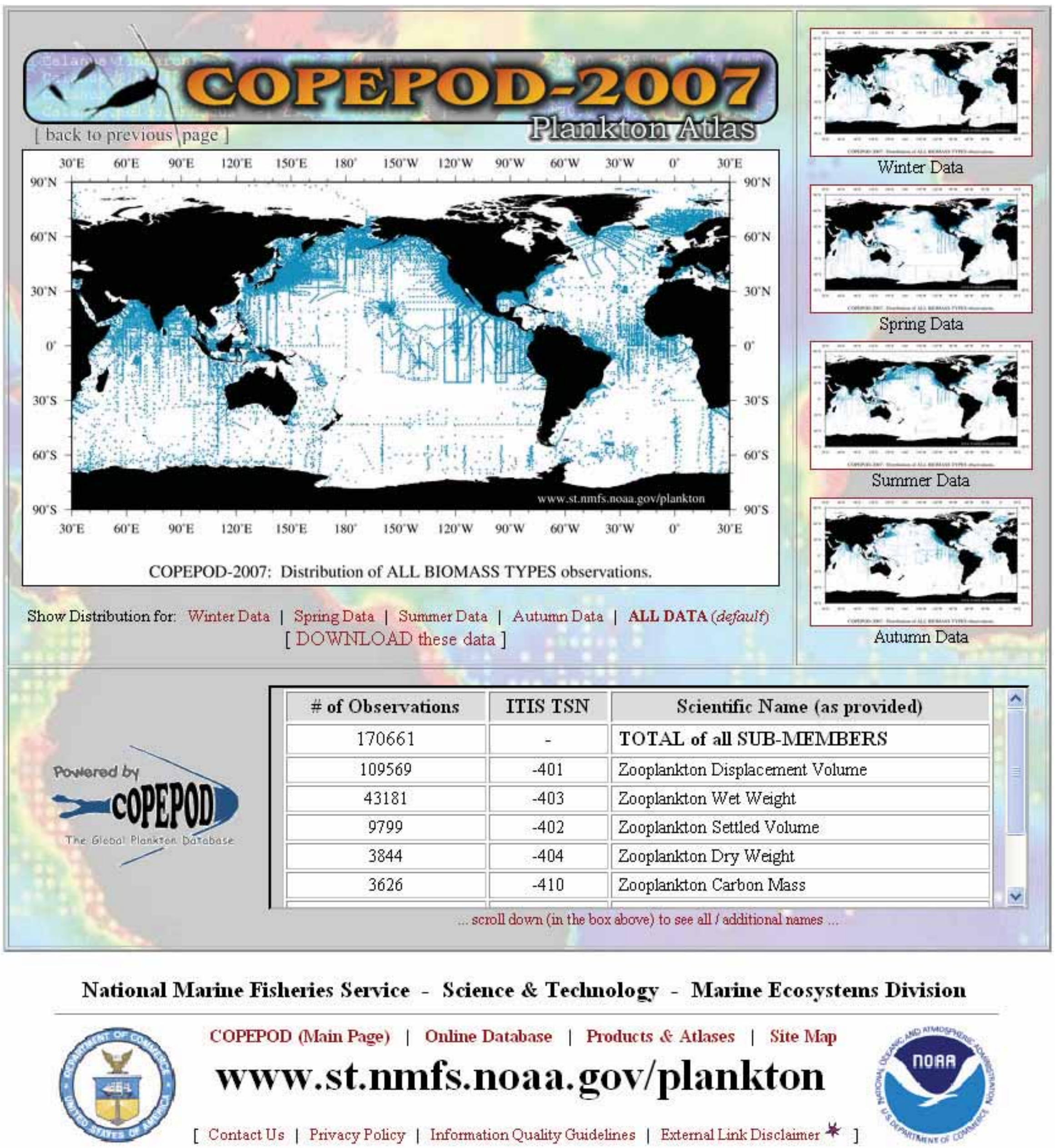


Figure 10: Example of the data content tables and plots found in the online Atlas section.

5.3 Global Zooplankton Biomass Fields

The global zooplankton biomass fields of *COPEPOD-2005* have been recreated to include new data from the Arctic, Antarctic, and South Pacific regions. These new fields were created using the same methods of *COPEPOD-2005* (O'Brien 2005), but now use the COPEPOD ranging system to detect, investigate, and/or exclude any data values flagged as >99.9% or <99.9% within their respective region and season. This additional step was found to help exclude very large values (typically caused by large phytoplankton blooms clogging the zooplankton net) and very low values (typically caused by gear failures).

The online *COPEPOD-2007* biomass fields include annual and seasonal mean carbon mass (mg-C/m^3), total wet mass (mg/m^3), total displacement volume (ml/m^3), total settled volume (ml/m^3), and total dry mass (mg/m^3).

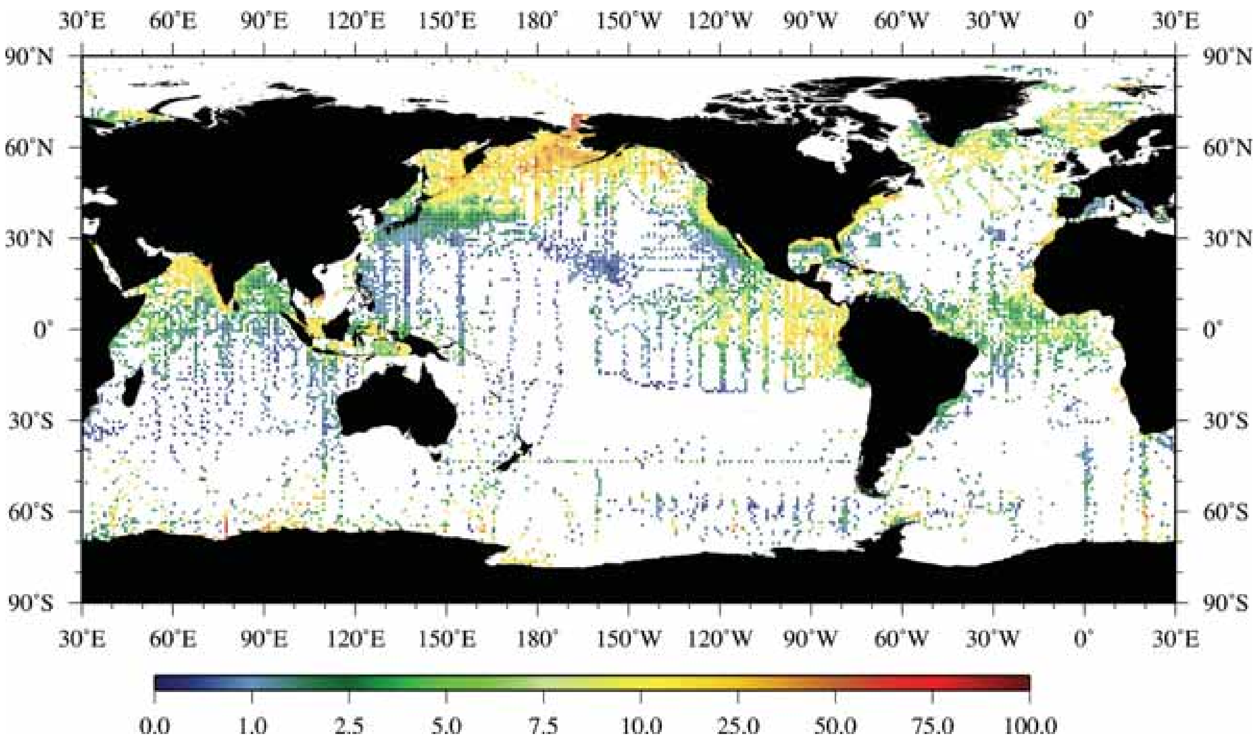


Figure 11: Annual mean zooplankton carbon mass (mg-C/m^3) fields from *COPEPOD-2007*.

REFERENCES

Conkright, M.E., T.D. O'Brien, L. Stathoplos, C. Stephens, T.P. Boyer 1998. World Ocean Database 1998, Volume 8: Temporal Distribution of Chlorophyll and Plankton Profiles. NOAA Atlas NESDIS 25, U.S. Gov. Printing Office, Wash., D.C., 129pp.

Conkright, M.E., H.E. Garcia, T.D. O'Brien, R.A. Locarnini, T.P. Boyer, C. Stephens, J.I. Antonov, 2002. World Ocean Atlas 2001, Volume 4: Nutrients. NOAA Atlas NESDIS 52, U.S. Government Printing Office, Wash., D.C., 392pp.

Harris, R.P, P.H. Wiebe, J. Lenz, H.R. Skjldal, and M. Huntley 2000. ICES Zooplankton Methodology Manual. Academic Press, 684 pp.

Margulis, L. and K.V. Schwartz 1998. Five Kingdoms: An Illustrated Guide to the Phyla of Life on Earth. W.H. Freeman & Company (New York), 520 pp.

O' Brien, T.D., M.E. Conkright, T.P. Boyer, D. Johnson, C. Stephens, R. Gelfeld 1998. World Ocean Database 1998 Volume 7: Temporal Distribution of Oxygen, pH, and Alkalinity Profiles. NOAA Atlas NESDIS 24, U.S. Gov. Printing Office, Wash., D.C., 235pp.

O'Brien, T.D., M. E. Conkright, T. P. Boyer, J. I. Antonov, O. K. Baranova, H. E. Garcia, R. Gelfeld, D. Johnson, R. A. Locarnini, P. P. Murphy, I. Smolyar, C. Stephens 2002. World Ocean Database 2001, Volume 7: Temporal Distribution of Chlorophyll and Plankton Data. NOAA Atlas NESDIS 48, U.S. Gov. Printing Office, Wash., D.C., 219pp.

O'Brien, T.D. 2005. COPEPOD: A Global Plankton Database. A review of the 2005 database contents and creation of new global zooplankton biomass fields. U.S. Dep. Commerce, NOAA Tech. Memo. NMFS-F/SPO-73, 136 p.

Listing of Investigators associated with the "*COPEPOD-2007*" data content.

[A] - T. ABE - Quaraisin ADNAN - A.P. ALEKSEEV - Arshad ALI - Moazzarn ALI – P.J. ALLEN - Usman ALWI - M. AMADEO - Shahid AMJAD - W.W. ANDERSON - Oscar ANGELES-REYES - Michel ANGOT - P. AOMORI - E. AOYAMA - Dharma ARIEF - O. H. ARINARDI - E. ASADA - Carin J. ASHJIAN - H. AUEL - Farooq AZAM - **[B]** - V. BAINBRIDGE - E.P. BALDINA - Richard T. BARBER - Richard A. BARKLEY - Clifford A. BARNES - Sonia BATTEN - Timothy R. BAUMGARTNER-McBRYDE – Allan BÉ - J. BEAUDOUIN - Robert L. BENWAY - T. BERNSHTEIN - Robert BIDIGARE - O.K. BILJEVA - George BOEHLERT - V. G. BOGOROV - Ph. BOURRET - E. BRATBERG - K.A. BRODSKII - Robert P. BROWN - E. BUSKEY - **[C]** - C. CABOCHE - Lisa CAMPBELL - H. CARADANT - David CARON - Louis CODISPOTI **-** D.V.P. CONWAY - R.T. COONEY - J. CORLETT - Charles S. COX - Maria M. CRIALES - Henryk CZKIETA - **[D]** - Michael DAGG - Donald S. DAY - K.M. DERIJUGIN - DJUMHAR - Maxwell S. DOTY - Hugh DUCKLOW - Janet DUFFY-ANDERSON - **[E]** - S. EBARA - Sayed Z. EL-SAYED - Bjornar ELLERTSEN - T. Saunders ENGLISH - **[F]** – Frank FERRARI - P. FOXTON - Mitsuo FUKUCHI - Y. FUKUDA - P. FUKUOKA - K. FURUHASHI - **[G]** - David GARRISON - Anne GAUZENS - J.W. GEHRINGER - R.S. GLOVER - G.P. GORBUNOV - Michel GOSSELIN - Julien R. GOULET - Marcia GOWING - E.H. GRAINGER - John R. GREEN - P.R.S. GREEN - Brian GRIFFITHS - J.R. GRINDLEY - Jean-Alfred GUEREDRAT - **[H]** - W. HAGEN - I. HALLGRIMSSON - H. HAMADA - Y. HAMADA - M. HANDA - Vagn HANSEN - H. HANSSEN – Jon HARE - P.M. HARGREAVES - H. HATTORI - Ken-Ichi HAYASHI - Erica HEAD - Thomas S. HIDA - Harry HIGGINS - Hans-Jurgen HIRCHE - Robert W. HOLMES - Rita HORNER - Takao HOSHIAI - L. HUTCHINGS - P. HYOUGO - **[I]** - A. G. ILAHUDE - Gennady V. ILYIN - M. IMAI - Yoshio INO - M. ISHII - Shingo ISHIKAWA - **[J]** - Luis Clemente JIMENEZ-PEREZ - P.L. JOHANSEN - Martin W. JOHNSON - Jack W. JOSSI - **[K]** - H. KAMAZAKI - M.M. KAMSHILOV - Hiroshi KANDA - Joe KANE - K. KAROHJI - Akiyoshi KATAOKA - Gerhard KATTNER - M. KAWACHI - S. KAWAE - Kouichi KAWAGUCHI - K. KAWASHIMA - KENDAL - V.L. KHMIZNIKOVA - I.A. KISELEV - Wojciech KITTEL - Anthony H. KNAPP - A.N. KOLESNIKOV - K. KONDOU - T. KONNO - A. KOTTHAUS - N. KUBO - T. KUBODERA - T. KURAMOCHI - K. KURODA - Leonid KUZNETSOV - **[L]** - Mike LANDRY - Peter V.Z. LANE - S.B. LANE - R.J. LARA - Viktor V. LARIONOV - Bertha E. LAVANIEGOS - Frederik LEATEMIA - Jacques LEBOURHIS - Bill LI - Alan LONGHURST - A.N. LUSHOV - Waldo K. LYON - **[M]** - David L. MACKAS - Enrique F. MADELLI - Pavel R. MAKAREVICH - Isak MALIK - T. MANABE - T.O. MANLEY - K.H. MANN - Ogur MARPAUNG - John MARRA - Osamu MATSUDA - Y. MATSUDA - M. MATSUZAKI - T. MAYAMA - John A. McGOWAN - Skip McKINNELL - Edison MEILALA - T. MINODA - Y. MIYAMOTO - MOESOLL - E. MORIYAMA - Robert A. MORRIS - V.K. MORYAKOVA - Shigeru MOTODA - Kojiro MOTOMURA - Muswerry MUCHTAR - Nicolai MUMM - Erik MUXAGATA - **[N]** - Yasushi NAKAJIMA - Eugene L. NAKAMURA - T. NAKAMURA - M. NANDA - V.A. NESMELOVA - R.N. NISHIMOTO - K. NOMURA - Anugerah NONTJI - **[O]** – Todd D. O'BRIEN - Kazuko ODATE - T. ODATE - Mikiho OGURI - M. OHNO - Hideaki OHTSUKA - Y. OHYAMA - Robert OLSON – Lana ONG - K. ONISHI - Jay E. O'REILLY - Y. OSAKA - **[P]** - Taisoo PARK - La PAY - Chairil PELHOUPESSY – Bill PETERSON - PHIFE - Sergey PIONTKOVSKI - Ewa PRESLER - I. Yu. PRUSOVA - V.P. PTOKHOV - **[Q]** - D. QUEIROZ - **[R]** - T.S.S. RAO - R. REPELIN - Christoph RICHTER - J. RIVATON - G.A. ROBINSON - Stephen J. ROMAINE - Mike ROMAN - ROUKHIY - Maria I. ROUKHIYAINEN - Jury RUDJAKOV - RUSMINTARD - **[S]** - H. SAKAI - Jennifer SALTZMAN - Daniel SAPULETE - Hiroshi SASAKI - N. SATO - Hiroo SATOH - M.S. SAVICH - Vladimir M. SAVINOV - R. SEIKAI - R.A. SELVAKUMAR - Kenneth SHERMAN - Alexandra A. SHMELEVA - Daniel E. SMITH - David SMITH - Sharon SMITH - Walker O. SMITH - Walter O. SMITH - Eddy SOFYAN - A.A. SOLOVJOVA - Grieg STEWARD - SUDIRDJ - M. SUGIMOTO - Yukio SUGIMURA - SUMARS - Vijayaraghavan SUMITRA - SUNARY - SUPRAPT - SUSEN - A. B. SUTOMO - SUYANT – Toru SUZUKI - [**T**] - Eiji TAKAHASHI - Kazutaka TAKAHASHI - I. Yu. TAMOIKIN - S. TAMURA - Akira TANIGUCHI - Atsushi TANIMURA - Persadaan TARIGAN - David TERHELL - Delphine THIBAULT - T.G. THOMPSON - Richard THOMSON - H. THUTHUI - U. TILMANN - Sergey F. TIMOFEEV - R. TSUDA - Arao TSURUTA - D.Th. TUHUMENA - **[U]** - Syunshiro UENO - Juanita URBAN-RICH - P.I. USATSCHEV - **[V]** - M. VANNUCCI - M.A. VIRKETIS - **[W]** - Brenda J. WADDELL - WAHY - K. D. WALDRON - Kenneth D. WALDRON - Kentaro WATANABE - Y. WATANUKI - Benny WATTIMENA - G. WEISSBERG - David W. WELCH - Pat WHEELER – Peter WIEBE - Karen WISHNER - Zbigniew WITEK - Sam WOUTHUYZEN - **[Y]** - Ken-ichi YAMAMOTO - H. YAMAZAKI - J. YANO - M. YASUDA - G. YOKOE - M. YOSHIDA - Salili Ali YUSUF - **[Z]** - M.M. ZABELINA - Angelina A. ZELIKMAN - Maria ZMIJEWSKA – Aleksey ZUYEV

COPEPOD

The Global Plankton Database