

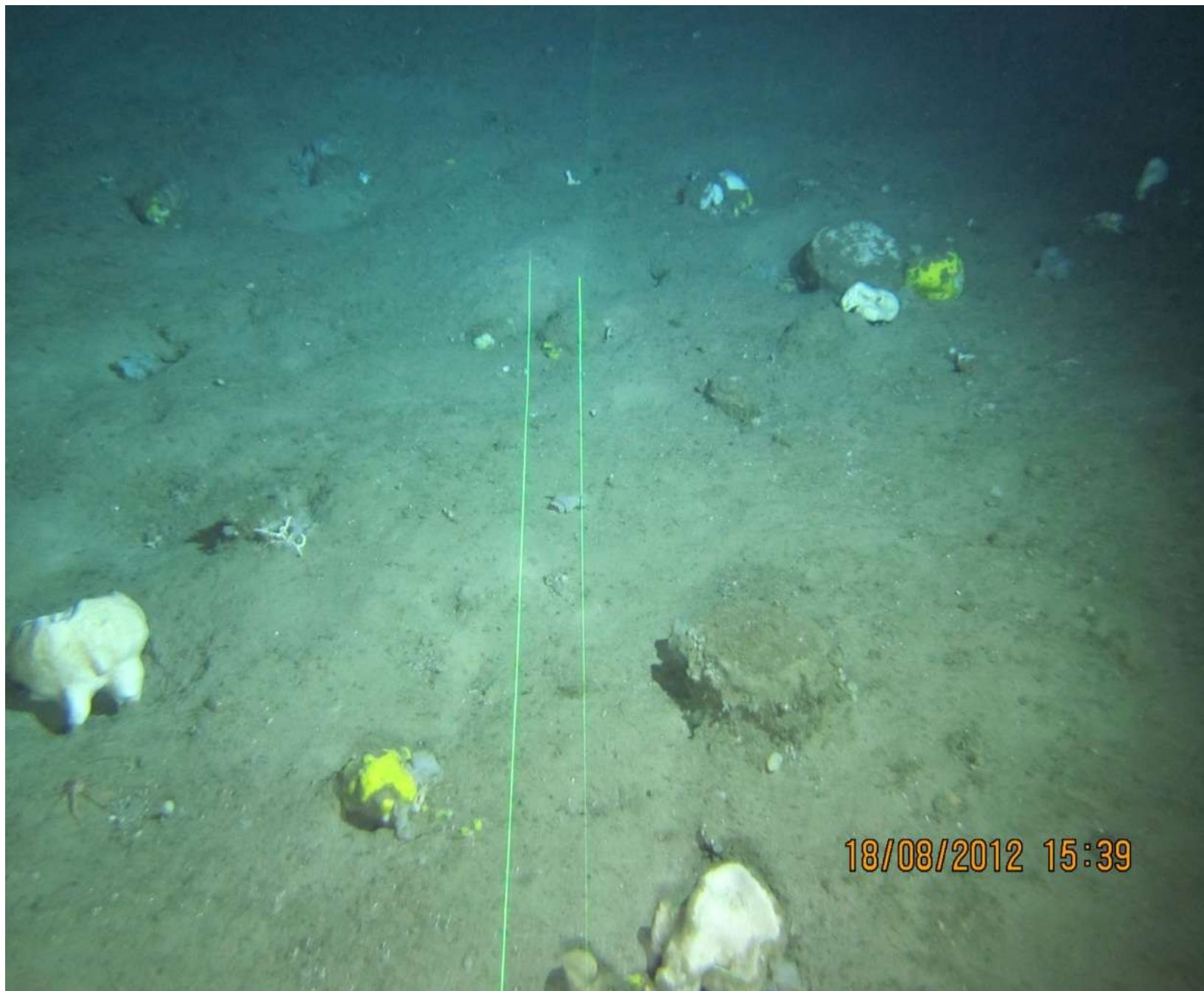
Web-based benthic databases: availability for science and quality issues



Sabine Cochrane (& friends)
Akvaplan-niva
Framsenteret, Tromsø

Outline

- Background
 - What are benthic data?
 - From paper to hard-disk to web
 - From reality to ideals and back again
- Quality issues
 - Names and authorities
 - Limitations
- Availability
 - Open access or restricted information
 - Publicising resources
- Conclusions
 - Optimal compromise
 - Relevance for MAREANO – for discussion!



Example: Benthic infauna

- Good indicators of sea floor conditions
- Quantitative sampling
- Sieve out fauna
- Identify and count
- Database and analysis
- Number taxa, individuals, diversity, similarity



Taxon lists – the building blocks

SPECIES NAME	1	2	3	4	5	6
Aglaophamus malmgreni	6	19	22	16	9	4
Abyssoninoe hibernica						1
Amage auricula						
Ampharete borealis				4		
Ampharete finmarchica			1		2	10
Ampharete goesi						2
Ampharete sp.						
Ampharetidae indet.						
Amphiteis gunneri			2	1		
Amphitrite groenlandica						
Anobothrus gracilis		3	8	7	2	14
Apistobranchus tullbergi			1	1		5
Aphelochaeta marioni	1	2	4	3	4	58
Aricidea hartmani						4
Aricidea nolani	7	2	8	5	3	18
Aricidea quadrilobata	3	3	6	4		
Artacama proboscidea						

Back to the background

- Paper archives
 - Increased focus re time series
 - Digitalising old species lists
- Hard disks and servers
 - From individuals to company servers
 - Property discussions....
 - Inter-company cooperation
 - National and international access
 - Example MOD database
- Web based
 - National
 - Usually themed
 - Wider focus
 - Larger scope – bigger undertaking



The process

- Data agreement
 - Authorship issues
- Everyone provides data
 - Needs standardising
 - Either work to get all to fill in schemes
 - Or leader does it for everyone
 - Once the names and positions are in
 - Positions need checking
 - Names need checking
 - Taxonomic levels need harmonising
- Several databases exist
 - Harmonising and standardising variable “stuff”
 - Challenges for user
 - Do all users know what they get??



FP7 NoE: Marine Biodiversity and Ecosystem Function



MarBEF Data System



[Introduction](#)

[Geography](#)

[Datasets](#)

[ERMS](#)

[EurOBIS](#)

A key task of the MarBEF Network is the integration of different resources related to marine biodiversity. The inventory of these resources can be found on this website. At the moment, this relational database includes information on different European marine biodiversity research sites and European marine biodiversity datasets. The European Register of Marine Species, ERMS and the European node of the Ocean Biogeographic Information System, EurOBIS is also accessible through this website. The terms of use of data are formulated in the [MarBEF data policy](#).

European Marine Biodiversity Research Sites

One of the objectives of the BIOMARE project was the selection of a network of Research Sites as the basis for long-term and large-scale marine biodiversity research in Europe. Among the 100 European Marine Biodiversity Research Sites that provide the geographical skeleton for the implementation of large-scale long-term

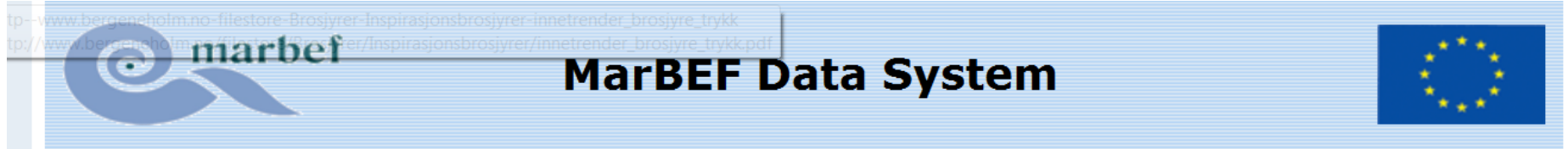


research in Europe, a small subset of Reference Sites has been selected. All the information about the BIOMARE sites is put into a fully searchable relational database and has a [geographical interface](#). Deep-sea, ocean pelagic, experimental or extreme habitat sites, proposed during the MarBEF project will be added to this database.

In order to link the different biodiversity datasets with a geographical area, a [geographical hierarchic gazetteer](#) was created. From each geographic placename, the related datasets are visible. The higher classification of the oceans and seas is based on the chart 'Limits of Oceans and Seas' 3rd edition (1953), published by the International Hydrogeographic Organisation (IHO).

European Marine Biodiversity Datasets

The MarBEF vision – and lasting legacy (?)



Introduction

Geography

Datasets

ERMS

EurOBIS

MarBEF Data - Terms of use

By downloading or consulting data from this website, the visitor acknowledges that he/she agrees with the [MarBEF data policy](#), and agrees to the following:

- If data are extracted from the MarBEF website for secondary analysis resulting in a publication, the MarBEF website should be cited. Online MarBEF databases should be cited as follows:
 - MarBEF (2004). European Marine Biodiversity Gazetteer. Available online at <http://www.marbef.org/data/gazetteer.php?p=browser>. Consulted on 2012-10-16.
 - MarBEF (2004). European Marine Biodiversity Research Sites. Available online at <http://www.marbef.org/data/sites.php>. Consulted on 2012-10-16.
 - MarBEF (2004). European Marine Biodiversity Datasets. Available online at <http://www.marbef.org/data/imis.php?module=dataset>. Consulted on 2012-10-16.
 - Costello, M.J.; Bouchet, P.; Boxshall, G.; Arvantidis, C.; Appeltans, W. (2008). European Register of Marine Species. <http://www.marbef.org/data/erms.php>. Consulted on 2012-10-16
[previously: Costello, M.J.; Bouchet, P.; Boxshall, G.; Emblow, C.; Vanden Berghe, E. (2004). European Register of Marine Species. Available online at <http://www.marbef.org/data/erms.php>. Consulted on date].
 - MarBEF (2004). European node of the Ocean Biogeographic Information System. Available online at <http://www.marbef.org/data/eurobis.php>. Consulted on 2012-10-16.
- If any individual datasource of MarBEF constitutes a significant proportion of the records used in the secondary analysis (e.g., more than 10% of the data are derived from this source), the individual data source should also be cited.
- If any individual datasource of MarBEF constitutes a substantial proportion of the records used in the secondary analysis (i.e. more than 25% of the data are derived from this source, or the data are essential to arrive at the conclusion of the analysis), the manager/custodian of this dataset should be contacted. It may be useful to contact the data source directly in case there are additional data that may strengthen the analysis or there are features of the data that are

From reality to ideals

Vol. 382: 221–224, 2009
doi: 10.3354/meps08045

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published April 30



THEME SECTION

Large-scale studies of the European benthos: the MacroBen database

Idea and Coordination: Paul J. Somerfield, Christos Arvanitidis, Edward Vanden Berghe

CONTENTS

Somerfield PJ, Arvanitidis C, Vanden Berghe E, Van Avesaath P, Hummel H, Heip CHR
MarBEF, databases, and the legacy of John Gray221–224
Vanden Berghe E, Claus S, Appeltans W, Faulwetter S, Arvanitidis C, Somerfield PJ and others
MacroBen integrated database on benthic invertebrates of

Arvanitidis C, Somerfield PJ, Rumohr H, Faulwetter S, Valavanis V, Vasileiadou A, Chatzigeorgiou G, Vanden Berghe E and others
Biological geography of the European seas: results from the MacroBen database265–278

And back to reality

Vol. 382: 239–252, 2009
doi: 10.3354/meps07963

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published April 30

Contribution to the Theme Section 'Large-scale studies of the European benthos: the MacroBen database'



Continental-scale patterns in benthic invertebrate diversity: insights from the MacroBen database

P. E. Renaud^{1,*}, T. J. Webb, A. Bjørgesæter, I. Karakassis, M. Kędra, M. A. Kendall, C. Labrune, N. Lampadariou, P. J. Somerfield, M. Włodarska-Kowalczyk, E. Vanden Berghe, S. Claus, I. F. Aleffi, J. M. Amouroux, K. H. Bryne, S. J. Cochrane, S. Dahle, S. Degraer, S. G. Denisenko, T. Deprez, C. Dounas, D. Fleischer, J. Gil, A. Grémare, U. Janas, A. S. Y. Mackie, R. Palerud, H. Rumohr, R. Sardá, J. Speybroeck, S. Taboada, G. Van Hoey, J. M. Węśławski, P. Whomersley, M. L. Zettler

¹Akvaplan-niva, Polar Environmental Centre, 9296 Tromsø, Norway

Comparison of biodiversity from 40 – 80 °N. Renaud et al (2009)

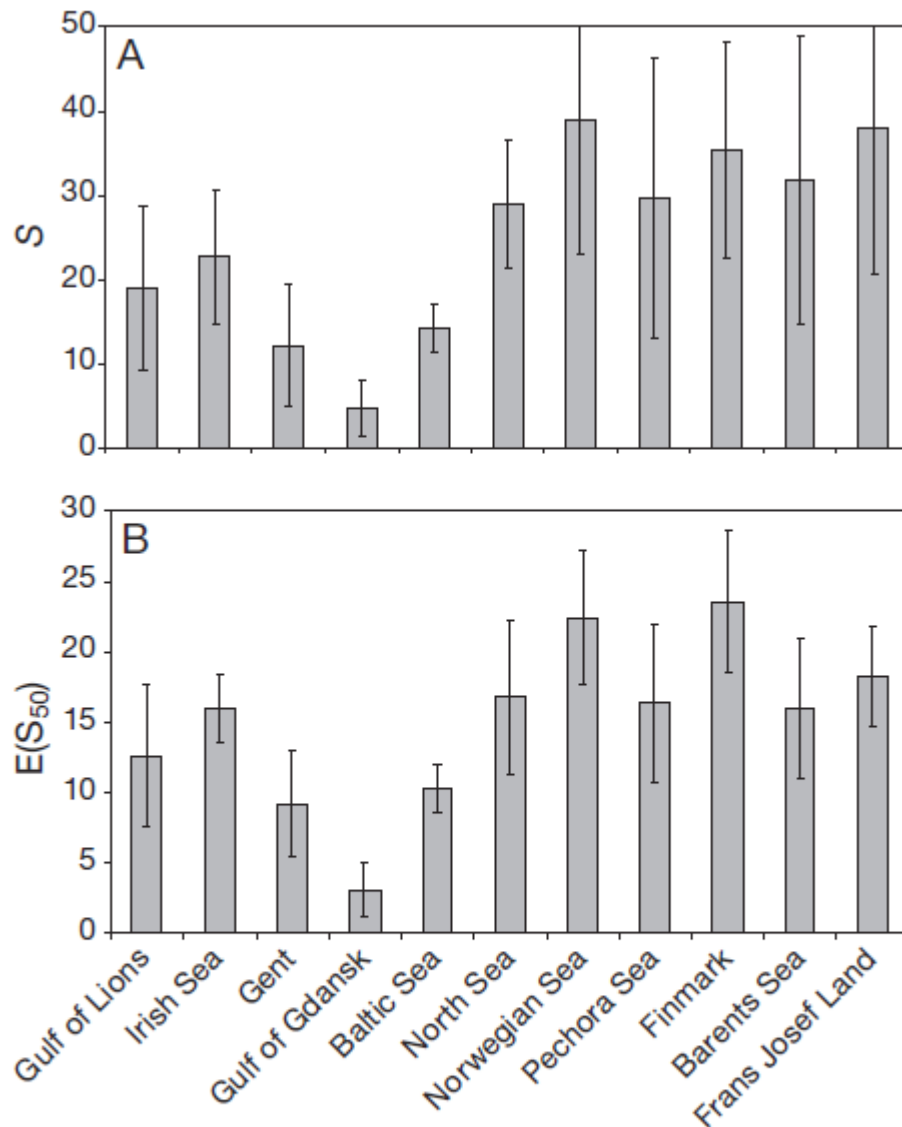
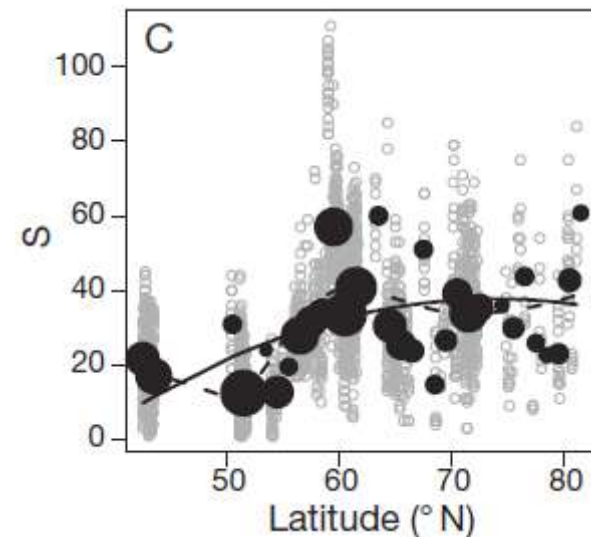
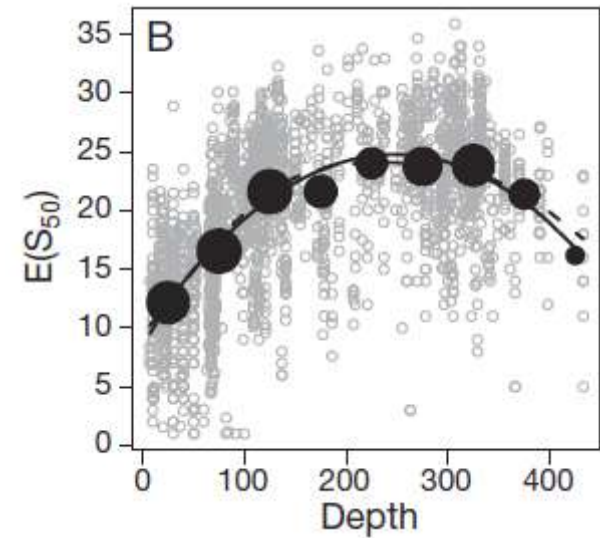


Fig. 3. (A) Total species richness, S , and (B) Hurlbert's expected number of species calculated for 50 individuals, $E(S_{50})$, in order of increasing latitude. Note low values for Baltic Sea datasets (Gulf of Gdansk, Baltic Sea). Error bars represent SD



Contribution to the Theme Section 'Large-scale studies of the European benthos: the MacroBen database'



Biological geography of the European seas: results from the MacroBen database

C. Arvanitidis^{1,*}, P. J. Somerfield, H. Rumohr, S. Faulwetter, V. Valavanis, A. Vasileiadou, G. Chatzigeorgiou, E. Vanden Berghe, J. Vanaverbeke, C. Labrune, A. Grémare, M. L. Zettler, M. Kędra, M. Włodarska-Kowalczyk, I. F. Aleffi, J. M. Amouroux, N. Anisimova, G. Bachelet, M. Büntzow, S. J. Cochrane, M. J. Costello, J. Craeymeersch, S. Dahle, S. Degraer, S. Denisenko, C. Dounas, G. Duineveld, C. Emblow, V. Escavarage, M. C. Fabri, D. Fleischer, J. S. Gray, C. H. R. Heip, M. Herrmann, H. Hummel, U. Janas, I. Karakassis, M. A. Kendall, P. Kingston, L. Kotwicki, J. Laudien, A. S. Y. Mackie, E. L. Nevrova, A. Occhipinti-Ambrogi, P. G. Oliver, F. Olsgard, R. Palerud, A. Petrov, E. Rachor, N. K. Revkov, A. Rose, R. Sardá, W. C. H. Sijm, J. Speybroeck, G. Van Hoey, M. Vincx, P. Whomersley, W. Willems, A. Zenetos

¹Institute of Marine Biology and Genetics, and Institute of Marine Biological Resources,
Hellenic Centre for Marine Research, PO Box 2214, Heraklion 71003, Greece

Data integration for European marine biodiversity research: creating a database on benthos and plankton to study large-scale patterns and long-term changes

**Leen Vandepitte · Bart Vanhoorne · Alexandra Kraberg · Natalie Anisimova ·
Chryssanthi Antoniadou · Rita Araújo · Inka Bartsch · Beatriz Beker · Lisandro Benedetti-Cecchi ·
Iacopo Bertocci · Sabine Cochrane · Keith Cooper · Johan Craeymeersch · Epaminondas Christou ·
Dennis J. Crisp · Salve Dahle · Marilyse de Boissier · Mario de Kluijver · Stanislav Denisenko ·
Doris De Vito · Gerard Duineveld · Vincent Escaravage · Dirk Fleischer · Simona Frascchetti ·
Adriana Giangrande · Carlo Heip · Herman Hummel · Urszula Janas · Rolf Karez ·
Monika Kedra · Paul Kingston · Ralph Kuhlenkamp · Maurice Libes · Peter Martens ·
Jan Mees · Nova Mieszkowska · Stella Mudrak · Ivka Munda · Sotiris Orfanidis ·
Martina Orlando-Bonaca · Rune Palerud · Eike Rachor · Katharina Reichert ·
Heye Rumohr · Doris Schiedek · Philipp Schubert · Wil C. H. Sistermans · Isabel Sousa Pinto ·
Alan J. Southward · Antonio Terlizzi · Evagelia Tsiaga · Justus E. E. van Beusekom ·
Edward Vanden Berghe · Jan Warzocha · Norbert Wasmund · Jan Marcin Weslawski ·
Claire Widdicombe · Maria Wlodarska-Kowalczyk · Michael L. Zettler**




Global Biodiversity Information System (GBIF)




GLOBAL
BIODIVERSITY
INFORMATION
FACILITY

[SPECIES](#) [COUNTRIES](#) [DATASETS](#) [OCCURRENCES](#) [SETTINGS](#)
[ABOUT](#)

```
<?xml version="1.0" encoding="UTF-8"
<response xmlns="http://rs.tdwg.org/t
<header>
<source accesspoint="http://145.18.162
<software name="TapirLink" version="0.2(re
```



... free and open access to biodiversity data




Search
species/country/dataset

Search

Welcome to the GBIF Data Portal

Access 388680911 data records (340362391 with coordinates) shared via the GBIF network.
To learn how to use this site, please see [About](#).
To tune this site for smaller displays, see [Settings](#).
[Version 1.2.6](#) - click here to see what is new!



Finding datasets



Explore Species

Find data for a species or other group of organisms.

Species

Information on species and other groups of plants, animals, fungi and micro-organisms, including species occurrence records, as well as classifications and scientific and common names.

Example species:

Puma concolor (Linnaeus, 1771)



Explore Countries

Find data on the species recorded in a particular country, territory or island.

Countries

Information on the species recorded in each country, including records shared by publishers from throughout the GBIF network.

See data for:

Norway



Explore Datasets

Find data from a data publisher, dataset or data network.

Datasets

Information on the data publishers, datasets and data networks that share data through GBIF, including summary information on 10067 datasets from 422 data publishers.

Latest dataset added:

Vietnam Type Culture Collection

Selecting data



This map only shows records with coordinates (**16065014** records from a total of **17267823** records). There are **15082416** occurrences within the viewed area. (21°W, 45°N, 59°W, 85°N)
Disclaimer: Maps depict density of data registered within the GBIF network index and not necessarily true species occurrence density gradients.

▶ [View all occurrences in Norway within the viewed area \(21°W, 45°N, 59°W, 85°N\)](#)

Resources providing data for Map

[show/hide](#)

Dataset	Count	Non-Georeferenced Count (does not appear on map)
(Appendix 6) Benthic foraminifera at termination I in sediment core GIK23256-1 in the Arctic Ocean PANGAEA - Publishing Network for Geoscientific and Environmental Data	253	0
(Appendix 6) Benthic foraminifera at termination I in sediment core PS1906-2 in the Greenland Sea PANGAEA - Publishing Network for Geoscientific and Environmental Data	663	0

Ocean Biogeographic Information System



Welcome to OBIS!

Last updated on Thu, 2011-01-13 09:49. Originally submitted by evberghe on 2010-05-25 15:58.

OBIS allows users to search marine species datasets from all of the world's oceans.



With our evolving OBIS database repository, users can identify biodiversity hotspots and large-scale ecological patterns, analyze dispersions of species over time and space, and plot species' locations with temperature, salinity, and depth.

To search the database, please select the "[Search Data](#)" option in the toolbar above.

SEARCH OBIS WEB PAGES (FOR DATA GO TO SEARCH DATA)

Search

LATEST NEW DATA LOAD, 27 JULY 2012

Follow us on Facebook, Twitter, LinkedIn, Mendeley and SlideShare.

USER LOGIN

Username: *

Password: *

Log in

«seek and thee shall find»

Dataset Search ?									
Search by name <input type="text" value="enter part of provider or dataset name"/>					Multi-select Select				
Provider name ▲	#datasets	#records		Dataset name ▲	ID	#taxa	#rec...	Years	
Argentinean RON	15	203,661	<div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> </div>	ChEssBase	1696	1,085	3,715	Could not...	
Australian Antarctic Data Centre	50	813,355		Eelgrass Community across an eutrop...	1630	63	231	2007 to 2..	
Australian Institute of Marine Scie...	21	239,711		HMAP-History of Marine Animal Popula...	7	680	313,587	1611 to 2..	
Bigelow Laboratory for Ocean Sci...	3	4,108		Natural Geography In Shore Areas (Na...	61	3,494	67,398	Could not...	
ColdWaterCorals	1	6,553		Pacific Ocean Shelf Tracking	64	11	243,728	2004 to 2..	
COMARGE	3	29,227		SeamountsOnline (Seamount Biota)	10	3,792	18,632	1869 to 2..	
CoML	10	788,842		Taxonomically comprehensive assess...	1456	546	130,399	1962 to 2..	
ECOCEAN_WhaleSharks	1	8,417		TOPP Fish (TOPP)	1670	1	9,593	2002 to 2..	
EurOBIS	257	11,100,673		TOPP northern elephant seal (Miroun...	605	1	1,445	Could not...	
FishBase	10	720,562		ZooGene A DNA Sequence Database...	11	43	114	1989 to 2..	
Gulf of Maine Census of Marine L...	2	1,216							
Hexacorals	1	64,518							
ICoMM	1	898,945							
IndOBIS	2	48,657							

From paper pile to electronic pile...

- Many parallel databases
- Some are linked, others not
- Quality disclaimers
- Which to choose?
 - Are they compatible??



Why might data be incompatible?

- Discovery of taxa new to science
 - Early descriptions very general
 - Previous single "species" can be split into several
 - Reassigned to other "genera"
- Different identifier practices
 - Expertise
 - Literature
 - Time pressure
- Sampling bias
 - Between institutes
 - between gears
 - sample size



Good girls don't...



Back to the building blocks

SPECIES NAME	1	2	3	4	5	6
Aglaophamus malmgreni	6	19	22	16	9	4
Abyssoninoe hibernica						1
Amage auricula						
Ampharete borealis				4		
Ampharete finmarchica			1		2	10
Ampharete goesi						2
Ampharete sp.						
Ampharetidae indet.						
Amphiteis gunneri			2	1		
Amphitrite groenlandica						
Anobothrus gracilis		3	8	7	2	14
Apistobranchnus tullbergi			1	1		5
Aphelochaeta marioni	1	2	4	3	4	58
Aricidea hartmani						4
Aricidea nolani	7	2	8	5	3	18
Aricidea quadrilobata	3	3	6	4		
Artacama proboscidea						

The problem



European Register of Marine Species



MarBEF Data System

[Introduction](#)[Geography](#)[Datasets](#)[ERMS](#)[EurOBIS](#)

ERMS Taxon search

(65,000 names; 33,000 accepted species; 2,000 References)

Enter the name of the taxon you want to look up. Genus and subgenus names should be included for species. Valid wildcards are '%' and '_' ('%' replaces zero or more characters, '_' replaces a single character; click [here](#) for details and examples). If you're not sure of the genus and/or subgenus of a species, replace them with a '%', followed by a space.

Search

e.g. Chromadora kreisi, Siriella, ...

☐ Limit to accepted taxa

☒ Limit to marine taxa

Taxon rank

Limit to
taxa belonging to
e.g. Mysidae. Only taxa with rank above genus will be returned.

Search

Name check function

Introduction	Geography	Datasets	ERMS	EurOBIS
--------------	-----------	----------	------	---------

ERMS Taxon match

You can use the ERMS Taxon Match Tool ([credits](#)) to automatically match your species list or taxon list with ERMS. After matching, the tool will return your file with the AphiaID's, valid names, authorities, ERMS classification and/or any other output you selected. [\[View manual\]](#)
For performance reasons, the limit is set to 1,500 rows. You can send larger files to info@marinespecies.org and we will return the results to you as soon as possible.

File*
Allowed filetypes: Plain text [TXT], Comma Separated [CSV] & Excel Sheet [XLS, XLSX]

Row delimiter ☐ First row contains column names

Column delimiter

Match authority ☐

Match upto Higher taxa only possible if a full classification is given in additional columns

Limit to taxa belonging to

Match against ☒ ERMS ☐ WoRMS

Output ☒ AphiaID ☐ LSID ☐ TSN ☒ ScientificName ☐ Authority ☒ Accepted name ☐ Classification
☐ Qualitystatus ☐ Taxon status ☐ Environment ☐ Citation

WoRMS Taxon list

47 matching records. Click on one of the taxon names listed below to check the details. [[new search](#)]

Euchone Malmgren, 1866
Euchone alicaudata Moore & Bush, 1904
Euchone analis (Kröyer, 1865)
Euchone arenae
Euchone bansei Ruff & Brown, 1989
Euchone barnardi Reish, 1968 accepted as *Euchone incolor*
Euchone capensis Day, 1961
Euchone chiade
Euchone cochranæ Fitzhugh, 2002
Euchone cortezi Reish, 1968 accepted as *Euchone arenae*
Euchone derjugini Uschakov, 1950
Euchone elegans Verrill, 1873
Euchone eniwetokensis Reish, 1968
Euchone hancocki Banse, 1970
Euchone heteroseta Hartman, 1978
Euchone incolor
Euchone lawrencii McIntosh, 1916
Euchone letterstedti (Kinberg, 1866) accepted as *Chone letterstedti* (Kinberg, 1866)
Euchone limicola [auctt.] accepted as *Euchone limnicola* Reish, 1959
Euchone limnicola Reish, 1959
Euchone longifissurata Uschakov, 1950
Euchone magna Moore, 1923 accepted as *Chone magna* (Moore, 1923)
Euchone magna (Fauchald, 1972)
Euchone normanni McIntosh, 1916 accepted as *Euchone rubrocincta normanni* (McIntosh, 1916)
Euchone olegi Zachs, 1933
Euchone pallida Ehlers, 1908
Euchone papillosa (Sars, 1851)
Euchone pararosea Giangrande & Licciano, 2006
Euchone perseyi (Zenkewitsch, 1925)
Euchone pseudolimnicola Giangrande & Licciano, 2006
Euchone purpurea Tauber, 1879
Euchone quadrisegmenta Zhao, Westheide & Wu, 1993
Euchone rosea Langerhans, 1884
Euchone rubella Ehlers, 1871 accepted as *Euchone analis* (Kröyer, 1865)

World Register of Marine Species



Home
About
Search taxa
Taxon tree
Literature
Distribution
Specimens
Match taxa
Editors
Statistics
..

Search WoRMS

Common name
contains
e.g. fish, whale

Scientific name
begins with
e.g. Delphinus delphis

[\[Advanced search\]](#)

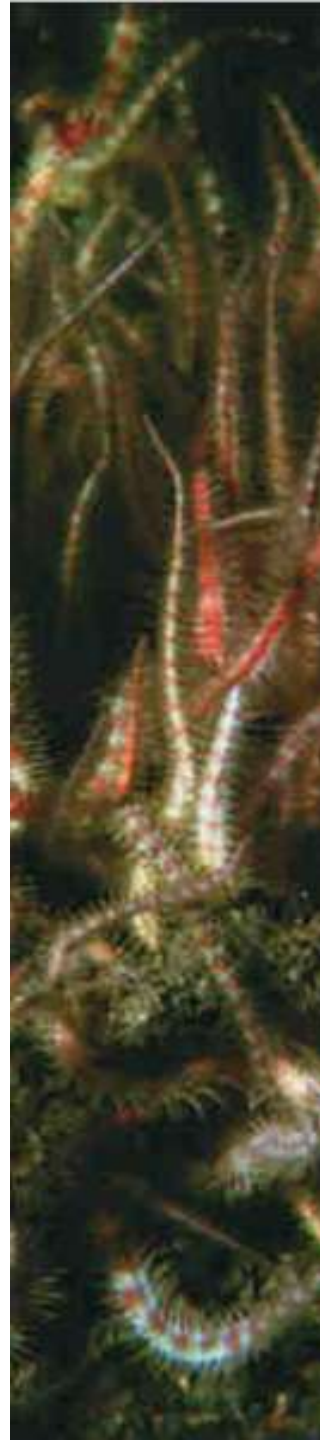
Problems solved and problems remaining

- Solved
 - Taxonomic synonyms
 - Spelling mistakes (Pectinaridae / Pectinariidae etc)
- Remaining «wriggles»
 - Identifier differences
 - several «correct» species
 - Different levels – the –sp. ghost
 - Historical data
 - Name changes without ability to trace which now is valid

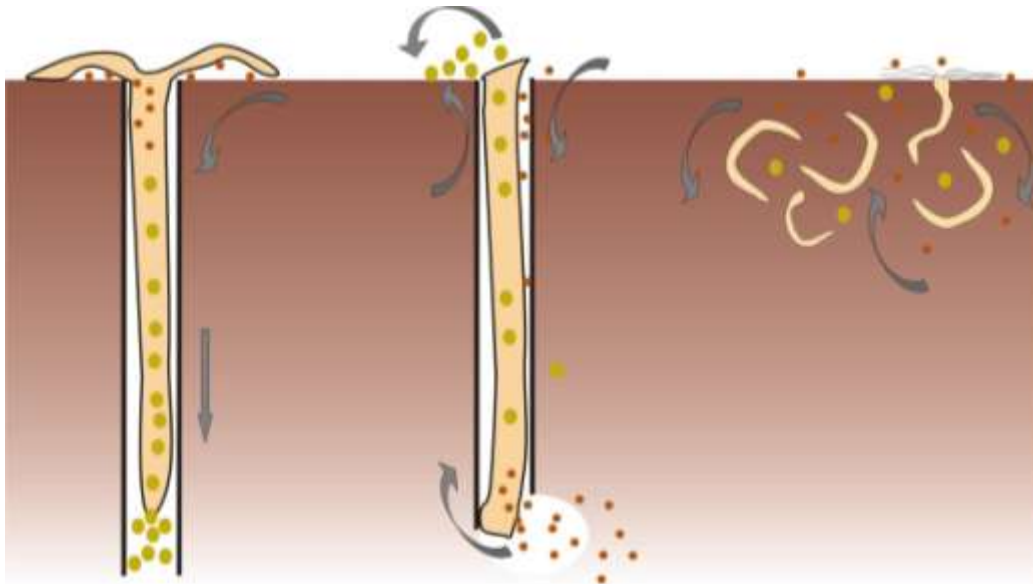


Possible approaches

- Surrogacy
 - Selected taxa
 - Ratios between taxa
- 'Lumping' to higher taxon
 - Genus?
 - Family?
 - Beware!!
 - Consider information content.....
 - Information value?
- Ecological function
 - Names used as shorthand for function info anyway!
 - Functional diversity has value in ecological studies
 - Assess long term change in terms of ecosystem function



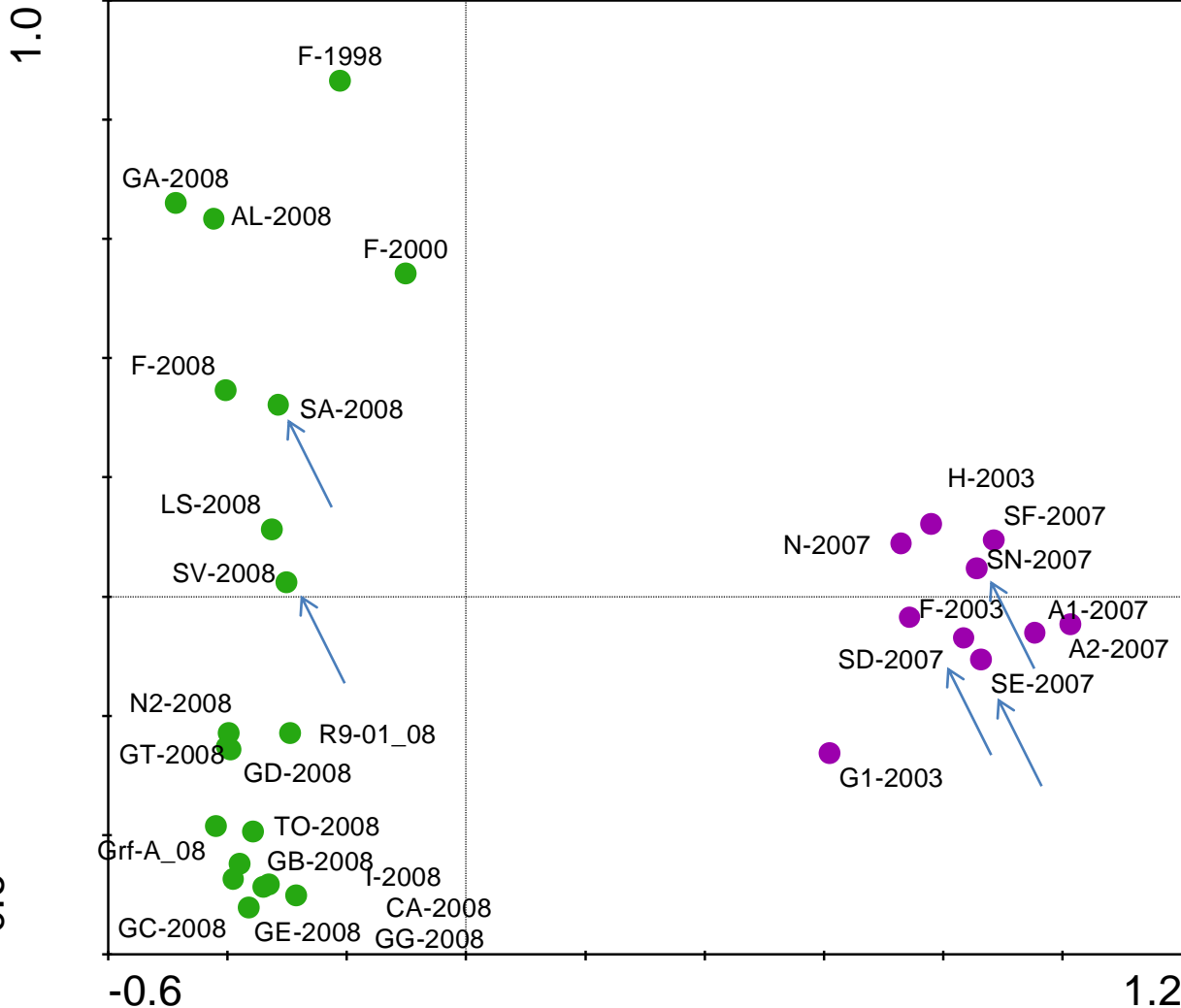
Function – what do benthic fauna do?



List of functional groups

	1	2	3	4	5	6
MF	22	22	142	29	140.5	156
SF	49.2	18	166	184	311	1296
MG	8	2	9	2	33	9
MSC	94	54	155	84	322	217
SFC	4	4	3	2	11	12
MDD	270	273	197	325	108	302
MSDC	36	33	95	34	22	38
SDD	516	42	300	166	24	97
MSD	194	131	298	907	742	581
SSD	33	4	45	52	52	88
MSF	3	5	26	16	95	102

Spatial patterns: names



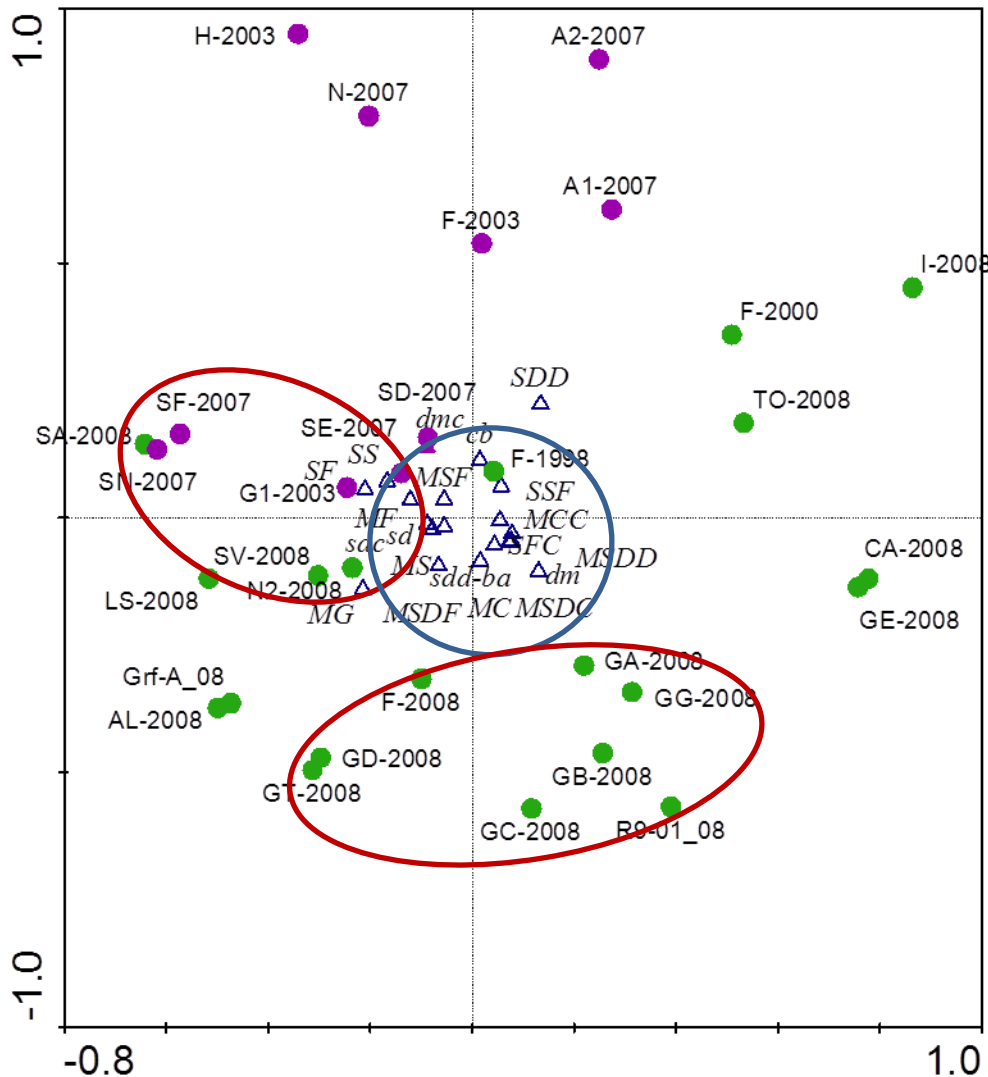
Dots represent 8-50 stations

Huge data resource

Institute bias??

How to extract ecological info from this?

Spatial patterns: functional groups



- Institute bias removed
- (different name = same function)
- Spread of stations
 - Sediment composition
 - Function of depth



A question of scale

-



-



Conclusions – implications for MAREANO

- Decide what is wanted
 - Max profiling of datasets
 - Control of publications?
 - Altruistic use of datasets?
- Decide on level of availability
 - Presence/absence or full data
 - Station/species maps, linked
- Compare with other datasets
 - Norway (MOD)
 - North Sea
 - Atlantic
 - Pan-European
- Who is going to do the harmonising?



Thank you!

