



MacroBen integrated database on benthic invertebrates of European continental shelves: a tool for large-scale analysis across Europe

E. Vanden Berghe^{1,2,*}, S. Claus, W. Appeltans, S. Faulwetter, C. Arvanitidis, P. J. Somerfield, I. F. Aleffi, J. M. Amouroux, N. Anisimova, G. Bachelet, S. J. Cochrane, M. J. Costello, J. Craeymeersch, S. Dahle, S. Degraer, S. Denisenko, C. Dounas, G. Duineveld, C. Emblow, V. Escaravage, M. C. Fabri, D. Fleischer, A. Grémare, M. Herrmann, H. Hummel, I. Karakassis, M. Kędra, M. A. Kendall, P. Kingston, L. Kotwicki, C. Labrune, J. Laudien, E. L. Nevrova, A. Occhipinti-Ambrogi, F. Olsgard, R. Palerud, A. Petrov, E. Rachor, N. Revkov, H. Rumohr, R. Sardá, W. C. H. Sistermans, J. Speybroeck, U. Janas, G. Van Hoey, M. Vincx, P. Whomersley, W. Willems, M. Włodarska-Kowalczyk, A. Zenetos, M. L. Zettler, C. H. R. Heip

¹Flanders Marine Institute, Wandelaarkaai 7, 8400 Oostende, Belgium

²Present address: IOBIS Secretariat, Institute of Marine and Coastal Sciences, Rutgers University, 71 Dudley Road, New Brunswick, New Jersey 08901, USA

ABSTRACT: We describe an integrated database on European macrobenthic fauna, developed within the framework of the European Network of Excellence MarBEF, and the data and data integration exercise that provided its content. A total of 44 datasets including 465 354 distribution records from soft-bottom macrobenthic species were uploaded into the relational MacroBen database, corresponding to 22 897 sampled stations from all European seas, and 7203 valid taxa. All taxonomic names were linked to the European Register of Marine Species, which was used as the taxonomic reference to standardise spelling and harmonise synonymy. An interface was created, allowing the user to explore, subselect, export and analyse the data by calculating different indices. Although the sampling techniques and intended use of the datasets varied tremendously, the integrated database proved to be robust, and an important tool for studying and understanding large-scale long-term distributions and abundances of marine benthic life. Crucial in the process was the willingness and the positive data-sharing attitude of the different data contributors. Development of a data policy that is highly aware of sensitivities and ownership issues of data providers was essential in the creation of this goodwill.

KEY WORDS: Macrofauna · Biogeography · Integrated database · Data management · Data rescue

—Resale or republication not permitted without written consent of the publisher—

INTRODUCTION

Often marine biological data are the result of projects with a limited temporal and spatial cover (Floen et al. 1993). Taken in isolation, datasets resulting from these projects are only of limited use in the interpretation of large-scale phenomena. More specifically, they fail to

be informative on a scale commensurate with the problems of global change that humankind is confronted with (Costello & Vanden Berghe 2006). Individual studies are restricted in the amount of data they can generate; but by combining the results from many studies, massive databases can be created that make analyses on a much-enhanced scale possible (Grassle 2000).

*Email: evberghe@iobis.org

Addresses for other authors are given in the Electronic Appendix at www.int-res.com/articles/suppl/m382p221_app.pdf

Such data have never been of greater importance considering the recent observation of major shifts of marine species due to global change. A recent report from the Marine Board of the European Science Foundation shows clear evidence of northward migrations of several marine-species populations from European seas (Philippart et al. 2007). Increasingly, vast thematic databases are being created, made possible by advances in computer technology. The Ocean Biogeographic Information System (OBIS), the data-integration component of the Census of Marine Life (CoML), assesses and explains the diversity and distribution of marine life through a network of linked databases (Grassle 2000). The Marine Biodiversity and Ecosystem Functioning (MarBEF) initiative includes implementation and networking of large-scale biodiversity research in Europe; part of this forms a contribution to the OBIS network through its European node, EurOBIS. Within the MarBEF working group Global Patterns of Marine Biodiversity across Ecosystems (Theme 1), an integrated database (MacroBen) on soft-bottom macrobenthos was compiled. The present paper describes the content, data integration, standardisation and functionalities of the MacroBen database.

CONSTITUENT DATASETS

Each dataset (Table 1) was archived and described in detail at the data centre of the Flanders Marine Institute (VLIZ). Describing datasets in a standardised way made it possible to create a searchable metadata inventory, thus facilitating data discovery and sharing. These metadata include information needed to decide on the relevance of a dataset in a particular context, e.g. where and when the data were collected, what the intended use of the dataset was, how data were collected, who played a role in the collection and the management of the data, where the data are stored and in what format and under which conditions they are available. All metadata descriptions are publicly available through the MarBEF website (www.marbef.org/data/dataset.php). Archiving the datasets prevents them from being lost by ensuring the long-term integrity of the data. The Integrated Marine Information System (IMIS), the metadata system maintained by VLIZ (Cattrijsse et al. 2006), was used as a tool for metadata management; this system is also used to manage information on behalf of the MarBEF network.

A unique citation, brief explanation of the original dataset objective and derived publications from each of the contributing datasets are listed in Appendix 1, together with a link to the extended metadata description available in IMIS. In Appendix 1, we also propose a standard dataset citation including title, dataset

collector(s) or custodian(s) (equivalent to author[s] or editor[s]), dataset owner (equivalent to publisher) and final year of sampling. We hope that this standard citation, or an equivalent one, will stimulate formal recognition of data sources in future re-use of the data.

DATA INTEGRATION

The MacroBen database integrates datasets on soft-bottom macrobenthic fauna from all European seas. A total of 44 datasets was collected, harmonised and integrated into a relational Access database. The integrated database includes both quantitative and qualitative data on organisms and several abiotic parameters ranging from the northeast Atlantic and the North Sea, Norwegian Sea and Arctic Ocean, the Baltic Sea and the Mediterranean and Black seas.

One of the main difficulties in integrating and comparing different datasets from various data providers is the harmonisation of the data; 3 types of harmonisation were performed: (1) taxonomic, (2) geographical, and (3) sampling methodology. All geographical coordinates were converted to decimal degrees. All datasets in which information on the datum was available used the WGS84 coordinate system. We assumed that this was also the case for those datasets where this information was missing.

Taxonomic names as recorded in the constituent datasets were matched with the European Register of Marine Species (ERMS). This authoritative taxonomic register provides a list of species occurring in the European marine environment, defined as up to the strand-line or splash zone above the high-tide mark and down to 0.5 salinity (Costello et al. 2001). Spelling mistakes were corrected, information on gender and life stage removed and stored in the appropriate fields, and the taxonomic name as recorded in the dataset was linked to the name as included as valid in the ERMS. Many names in the constituent datasets proved to be omissions in the ERMS, and were, in consultation with the ERMS taxonomic editors, subsequently added to the ERMS. A further description of the ERMS can be found in Cuvelier et al. (2006), and on the MarBEF website (www.marbef.org/data/erms).

Various sampling tools and methods were used over the different datasets. About 50% of the data was derived from sampling with Van Veen grabs. To a lesser extent the data were derived from sampling with a Smith-McIntyre, Ponar or APN grab, cores or other hand-operated devices. Two datasets provided data from museum collections. Overall, 33% of the data were qualitative in nature, while 30% contained not only specimen counts but also biomass data (Table 2). The sampling area from the quantitative datasets

Table 1. Datasets included in the MacroBen database with 2-letter station code. The ID refers to the metadata of the datasets stored in the Register of Resources

ID	Code	Dataset	Geographical area	Total no. of records	Total no. of stations	Temporal cover
North Sea and northeast Atlantic Ocean						
345	bi	BioMar survey of Ireland	Irish coastal zone	86 359	2397	1972–1996
1600	o5	BIOMÖR 1 dataset	Southern Irish, Celtic seas	5166	51	1989–1991
615	np	National Marine Monitoring Programme	North Sea, English Channel Celtic Sea	1161	40	2002–2003
534	pl	Plymouth Sound dataset	Plymouth Sound, English Channel	1343	44	1995
597	ba	Benthos Gironde Estuary	Gironde Estuary, Biscay	3019	13	1979–2004
599	ni	BIS dataset of SW Netherlands (1985–2004)	Southern part of Dutch continen- tal shelf, North Sea	136 677	15 564	1974–2005
633	ug	MacroDat Belgium	Belgian part of North Sea	33 995	1801	1977–2003
67	ns	North Sea Benthos Survey	North Sea	11 820	231	1985–1986
999	o3	Offshore reference stations, North/ Norwegian seas	North Sea	9904	30	1990–2002
Norwegian Sea and Arctic Ocean						
997	o4	Offshore reference stations, Norwegian/ Barents seas	North Sea, Norwegian Sea, Barents Sea	59 710	184	1991–2001
243	hs	Polish Arctic Marine Programme	Spitsbergen, Greenland Sea	603	34	2002
614	ko	Soft bottom community structure and diversity in Arctic Kongsfjorden	Spitsbergen, Greenland Sea	210	6	2003
533	ar	Arctic soft-sediment macrobenthos	Barents Sea and Arctic Ocean	1005	22	1991
998	o6	Offshore reference stations, Finnmark	Norwegian, Barents seas	11 898	53	1998–2000
12	o2	Northern Barents Sea 1992	Barents Sea	1919	10	1992
11	o8	Benthic fauna around Franz Josef Land	Barents Sea	2630	9	1992
461	o7	Benthic fauna around Pechora Sea	Barents Sea	2964	15	1992
Skagerrak and Baltic Sea						
601	ze	Macrozoobenthos Baltic Sea (1980–2005) as part of IOW Monitoring	Baltic Sea	3589	19	1980–2005
611	pu	Bay of Puck dataset	Baltic Sea	539	29	1996
612	gd	Gulf of Gdansk dataset	Baltic Sea	87	13	1995–2002
617	n3	N3 data of Kiel Bay	Baltic Sea	8944	1	1986–2004
618	70	Kiel Bay intercalibration dataset	Baltic Sea	1144	1	1995
11	of	Oslofjord 93–94	Skagerrak	1469	34	1993–1994
618	of	Oslofjord 96–97	Skagerrak	1096	23	1996–1997
Mediterranean and Black Sea						
603	bl	Soft Bottom of the Bay of Blanes	Balearic Sea	3754	2	1992–1997
213	gr	Redit dataset	Western Mediterranean	5642	92	1998
597	fb	Fauna Bentonica dataset	Adriatic Sea	2418	380	1934–1936
602	lm	LBMRev dataset	Adriatic Sea	9321	28	1985–2004
600	oc	Cesenatico dataset	Adriatic Sea	963	1	1996–2002
630	o1	Northern Aegean dataset	Aegean, Ionian seas	7277	419	1937
595	ap	Aegean polychaetes dataset	Aegean Sea	2215	70	1937–2000
613	ka	Cretan shelf macrofauna and macro- fauna of Greek polluted sites	Aegean, Ionian seas	11 861	199	1986–1997
598	do	Mop Crete dataset	Aegean Sea	2670	56	1988
622	M2	Aegean Sea coastal benthic communities, Geras Gulf (Mytilini)	Aegean Sea	4634	9	1986–1988
289	M3	Seasonal zoobenthos, Saronikos	Aegean Sea	2050	6	1989–1990
623	M6	Kalamitsi dataset	Ionian Sea	2037	22	1990–1991
624	M7	Kerkyra dataset	Ionian Sea	1015	12	1991–1992
625	M8	Zoobenthos Kyklades, Aegean Sea dataset	Aegean Sea	882	14	1986
272	M0	Gialova 1	Aegean, Ionian seas	1216	7	1995–1996
626	M1	MegFeod-Black Sea dataset – IBSS, Sevastopol	Black Sea	169	24	1986
627	M4	Strelbay-Black Sea dataset – IBSS, Sevastopol	Black Sea	297	18	1993
628	M5	Yalta-Black Sea dataset – IBSS, Sevastopol	Black Sea	308	26	1979–1986
629	M9	Laspibay-Black Sea dataset – IBSS, Sevastopol	Black Sea	689	33	1996
490	bo	Biocean dataset	Atlantic, Arctic, Baltic, Mediterranean seas	18 685	1071	1967–2001

ranged from 0.008 to 1.2 m² with 0.1 m² being the most common. In 23 % of the distribution records, replicate samples were taken.

CONCEPTUAL MODEL

The structure and concept of the MacroBen database and data management was based on an earlier system, created for a similar initiative: the North Sea Benthos Project 2000 (NSBP), an activity of ICES (Rees et al. 2007). In the NSBP initiative, data from the North Sea from various sources was brought together and analysed jointly. The data management for the NSBP was described in Vanden Berghe et al. (2007). The data management of both NSBP and MacroBen is based on a series of Access databases: 1 for each constituent dataset, and 1 to contain the integrated data. The latter also contains tools and a simple user interface to work with the data. At any point, the integrated data could be recompiled from the individual databases. This system made it possible to manage the data from individual datasets separately, and go through an iterative process of data cleaning and harmonisation. The structure of the Access databases containing the individual datasets depended on the formats employed by the data providers, and will not be discussed further here. Each of these individual databases contained a query that wrote a standard table that was then available for the integrator Access file to read.

The basic structure of the integrator database (Fig. 1) is based on a relatively limited number of tables: distribution records, taxonomic information, georeferenced sampling stations, abiotic data and metadata. Since the data were not edited in this database, no attempt was made at complete normalisation; the normalisation was done at the level of the Access databases containing the constituent datasets. Before analysis, the data structure was even further simplified, to a single flat table. This way, users were not forced to study the structure of the database before analysing the data. It also made export of the data to other programs such as spreadsheets trivial.

Both qualitative and quantitative distribution records were integrated into the distribution record (dr) table. This table stores the station name and a unique station code, code of the replicate sample, latitude and longitude, date, depth, surface area sampled, sampling depth, sediment type, mesh size, count and biomass, life stage, the original dataset code (db), a code for the standardised species name (aphiaID) and a free note. Species name, latitude and longitude, and a description of the origin of the data were the only mandatory fields; but extra information was available for many of the constituent datasets.

Table 2. Sampled area and equipment used over the different datasets. na = not applicable

Code	Area (m ²)	Equipment
bi	na (qualitative data)	Direct observation on seashores and by SCUBA divers
o5	0.1	Van Veen grab
np	0.1	Day grab
pl	0.008	Diver-operated cores
ba	Variable	Several (hand-corer [intertidal] and Smith-McIntyre grab [subtidal], suction grab and quantitative Sanders sledge)
ni	Variable	Several (cores, box-corer, Van Veen grabs, Flushing sampler)
ug	0.1026 and 1.2	Van Veen grab
ns	Variable	Van Veen grab
o3	0.1	Modified Van Veen grab (APN grab, patent Akvaplan-niva)
o4	0.1	Modified Van Veen grab (APN grab, patent Akvaplan-niva)
hs	0.1	Van Veen grab
ko	0.2	Diver-operated airlift system/corer
ar	0.1	Box-corer
o6	0.1	Modified Van Veen grab (APN grab, patent Akvaplan-niva)
o2	0.1	Van Veen grab
o8	0.1	Van Veen grab
o7	0.1	Van Veen grab
ze	0.1	Van Veen grab, dredge with 5 mm mesh size (qualitative)
pu	0.01	Eckman grab
gd	0.1	Van Veen grab
n3	0.1	Van Veen grab
70	0.1	Van Veen grab
of	0.1	Van Veen grab
of	0.1	Van Veen grab
bl	0.06	Van Veen grab
gr	0.1	Van Veen grab
fb	0.1 and 0.2	Petersen dredge
lm	0.1	Van Veen grab
oc	0.24	Van Veen grab
o1	Variable	Variable (museum collection)
ap	Variable	Variable (museum collection)
ka	0.1 and 0.035	Smith-McIntyre grab/core
do	0.1	Smith-McIntyre grab
M2	0.045	Van Veen grab, hand-operated
M3	0.05	Ponar grab
M6	0.05 and 0.1	Several (Ponar grab, Smith-McIntyre and Veen grab)
M7	0.2	Van Veen grab, hand-operated
M8	0.1	Smith-McIntyre sampler, hand-operated
M0	0.05	Van Veen grab, hand-operated
M1	0.25	'Ocean' grab
M4	0.08	Van Veen grab, hand-operated
M5	0.25	'Ocean' grab
M9	0.04	Van Veen grab
bo	na (qualitative data)	Variable

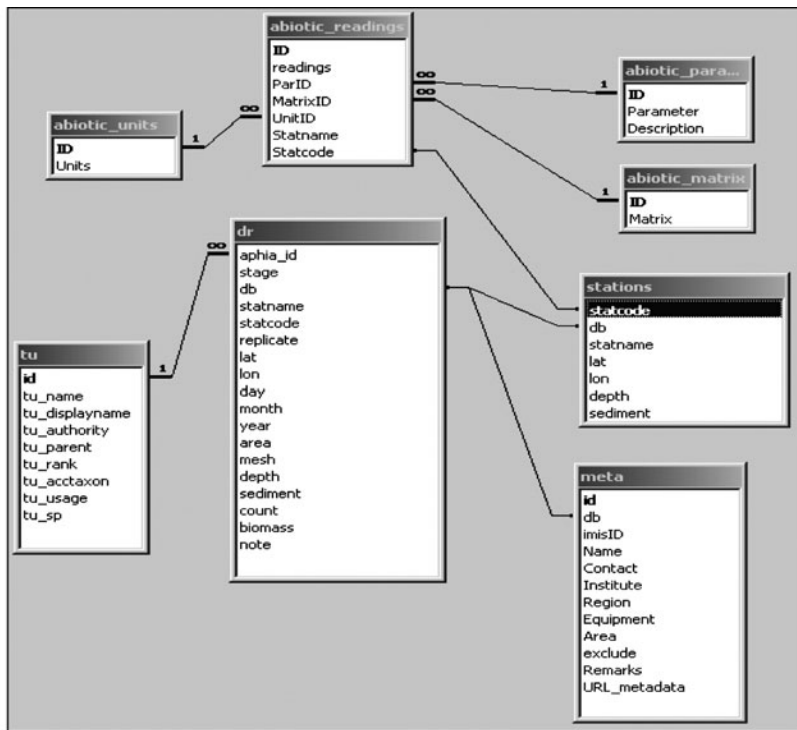


Fig. 1. Relational structure of the MacroBen database. Most table names and field names are self-explanatory; dr - distribution records; tu - taxonomic units; id - unique identifier for a record in that table. Lines show the relationships between the tables, with '1' and infinity signs indicating that one record in the first table can have many related records in the second table (e.g. one taxonomic unit has many distribution records)

A snapshot of ERMS was included in the MacroBen database, in the taxonomic unit (tu) table. The classification that is included in ERMS was used to create mechanisms to combine taxa into taxa of higher rank and to calculate several taxonomic variance measurements (Vanden Berghe 2006).

Abiotic variables measured at the stations, such as temperature, salinity and sediment type, were standardised and are stored in the table *abiotic_readings*. The measured parameters, the matrix and the units are stored in the tables *abiotic_parameters*, *abiotic_matrix* and *abiotic_units*.

The descriptive metadata of the different datasets are stored in the table called *meta* and linked with the biogeographical information through the unique dataset code (db). In the *meta* table, information was recorded that was essential for the interpretation of the data, especially information on sample sizes and methodology. Table 1 gives an overview of the datasets included in MacroBen, and was created on the basis of a report extracted from IMIS. Unique identifiers for each of the datasets are listed in the table, and can be used to find the relevant record in IMIS.

The database interface can be made available upon request, for similar initiatives, or to re-utilise some of

the coding to calculate the biodiversity measures. The Manuela Project, part of MarBEF, made use of the same database structure and tools (Vandepitte et al. 2007). Future work within Theme 1 of MarBEF will also make use of the system.

FUNCTIONALITIES

An interface was built on the integrated database, allowing users to manipulate data prior to analysis and allowing subselection of data at the level of dataset, species, or individual distribution records (Fig. 2).

The MacroBen database stores a set of built-in functions for calculation of taxonomic and diversity indices. Various options can be chosen for subselecting part of the data for analysis, such as the inclusion or exclusion of several datasets, selecting only the qualitative or quantitative data, only observations on adult specimens, restricting the data to a single taxonomic group or an identification rank or excluding manually non-macrobenic taxa. Temporal or spatial selections can

also be made. By indicating the proportion of samples in which a taxon has to be present and the minimum number of individuals in a sample to force the species to be present, rare species can be excluded from the selection. Excluding rare species has the effect of reducing the data matrix to a more manageable size (depending on the analytical objectives); it also minimises the effects of rare, possibly vagrant species, on the analysis (Gray 1981). Though the facility to conduct such filtering was considered important, the use of this feature is not necessarily recommended: modern computers and software are able to handle large matrices; a suitable choice of algorithms should make sure that the analysis demonstrates main effects, rather than those resulting from chance encounters of rare species.

The interface facilitates pooling of records using different criteria: replicates, taxonomy and life stages. For taxonomy, there were different lumping strategies: species level, genus level, family level, valid taxon or original name. The species-level lumping strategy keeps the highest taxonomic precision but sample size will decrease, since all incomplete identifications will be discarded. Aggregating records at genus or family level will decrease the taxonomic precision but will allow the inclusion of records only identified to genus or family,

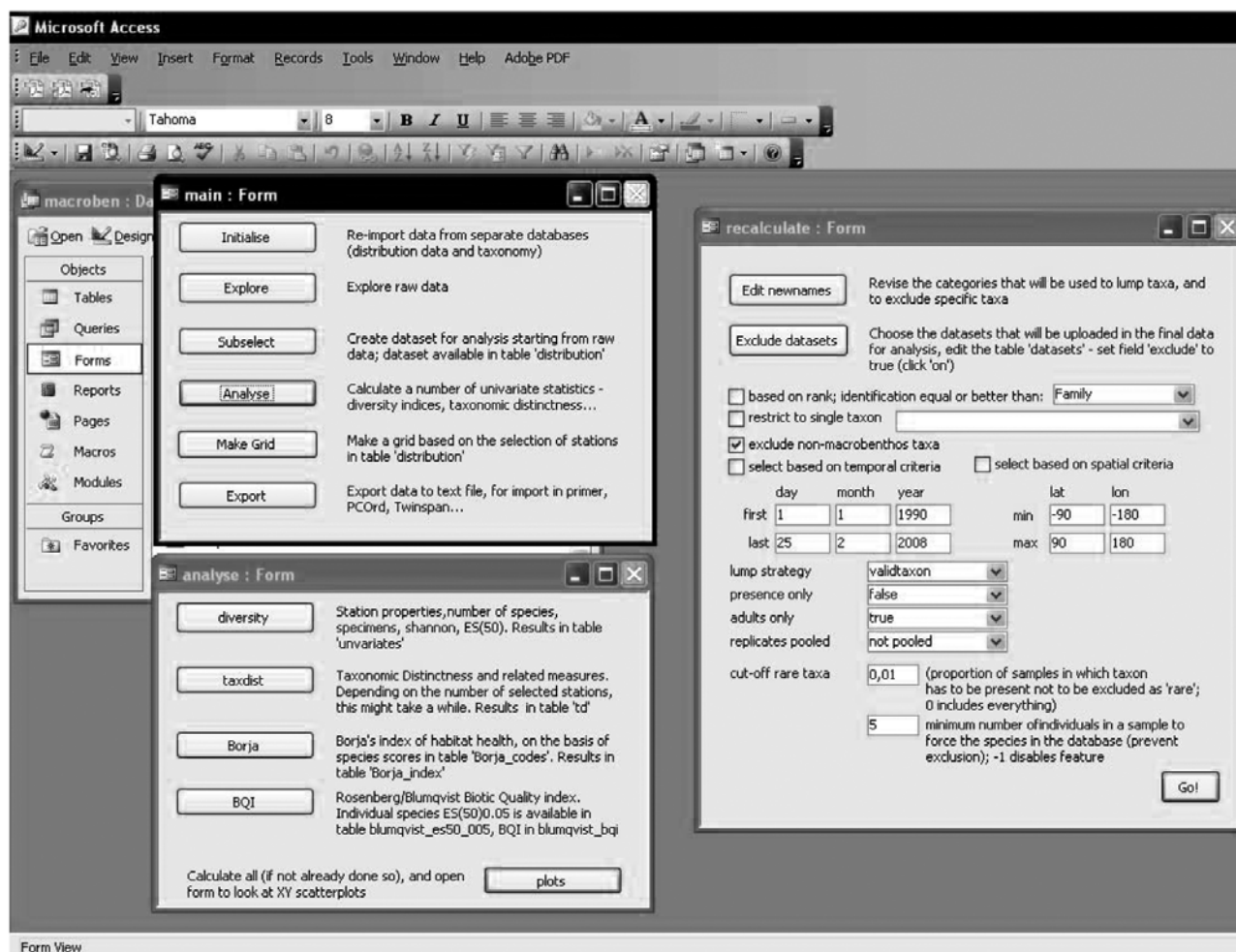


Fig. 2. User interface of the MacroBen database

respectively, thus increasing the number of records available for analysis. The main reason to include the taxonomic lumping was to increase the robustness of the conclusions based on the analysis. Aggregating to a higher taxonomic level has the advantage that resulting analyses become more robust, and less dependent on the potential problems posed by differences in identifications by different research groups.

After selection of the data for inclusion and pooling strategies, an intermediate table with density values (ind. m^{-2}) will be created, from which a variety of frequently used diversity and taxonomic indices can be calculated such as: (1) the Shannon-Wiener diversity index (Magurran 1998), (2) the Hurlbert diversity index calculated for 50 individuals, $E(S_{50})$ (Hurlbert 1971), (3) the Biotic Quality Index (Rosenberg et al. 2004), (4) AMBI indices of habitat health (Borja et al. 2000, 2003) and (5) taxonomic distinctness coefficients and related measures (Hurlbert 1971, Clarke & Warwick 1998, 1999, 2001). A graphic interface of the database allows 2-dimensional plotting of the calculated indices.

The different stations used in the selection can be projected (converted from latitude and longitude to distance in m north and east from a reference point) using the Lambert Azimuthal Equal Area or the Azimuthal Equidistant projection. This allows calculating distances between positions of stations, and aggregating data in grid cells, with the size of the cells controlled by the user.

After selection, calculation of the different indices and projection, the MacroBen database allows export of the species $\times$ Station matrix to a delimited text file, which can be imported into additional statistical analysis software, such as TWINSpan (Hill 1979) or PRIMER (Clarke & Warwick 2001). Both density and count data can be exported; densities are scaled to 1 m^2 and counts are the actual number of specimens counted. Possible export file types are in condensed format (.con), which can be imported in the TWINSpan package, an ASCII tab-delimited list (e.g. PCOrd) or an ASCII tab-delimited table (e.g. PRIMER). Another possible export is the taxonomic hierarchy, which can be used as an aggregation file in PRIMER.

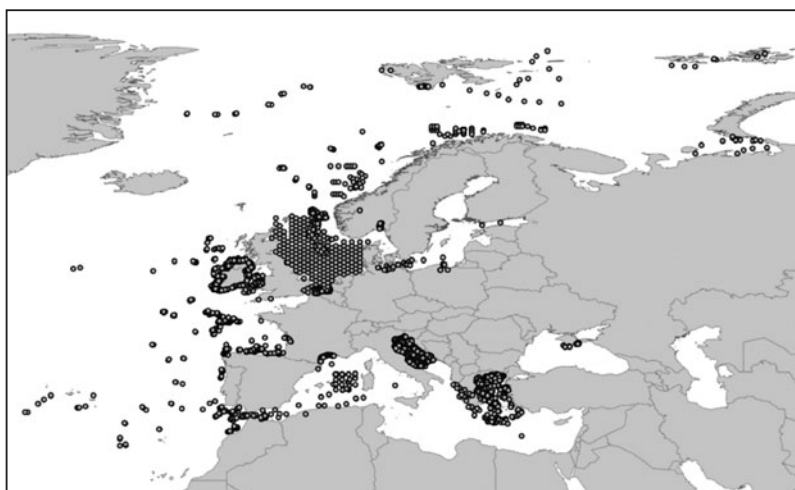


Fig. 3. Distribution map of the sampling stations of the integrated MacroBen database

CONTENT OF THE DATABASE

The MacroBen database contains data from 1937 to 2005 from different research and monitoring programmes and includes 22 897 sampled stations and 465 354 distribution records (Table 1). Depth is available for 416 312 distribution records (86%). There are 7481 taxa, of which 7203 were valid. The database contains 40 766 abiotic readings from 75 different parameters from both the sediment and water; 46 datasets provided by 24 European institutes were collected and integrated into the MacroBen database. The samples range geographically from 32° 2' 48" N to 81° 27' N and from 42° 55' 1" W to 58° 56' 35" E (Fig. 3).

The North Sea and northeast Atlantic region has the highest sampling density. Datasets from the Belgian part of the North Sea and the southern part of the Dutch continental shelf account for 37% of the total biotic records, while <0.05% of the total records in MacroBen are from the Skagerrak and Baltic Sea region. It is essential to allow for these strongly unequal sampling densities across the different geographical regions in evaluating the ecological basis for any observed patterns.

DATA POLICY

All data stored in the MacroBen database are subject to the declaration of mutual understanding for data sharing within MarBEF Theme 1. The policy implies that the participating institutes, organisations and/or the collector of the dataset remain owners of their contributed dataset, and in control of the conditions under which datasets can be used by a third party. Metadata,

as an essential tool for data discovery, is publicly available for all datasets, and includes the conditions of use. A description of the process that led to the policy, as well as a more complete description of the policy itself, is given in the introduction to this Theme Section. (Somerfield et al. 2008).

CONCLUSION

Probably the most challenging aspect in creating an integrated database is to convince data owners or custodians to contribute data. A relationship based on trust between data manager and data owner is therefore essential. Through networking activities (MarBEF Network of Excellence) at the European scale, an unambiguous data policy stating the restrictions and future use of data was drafted. Part of the discussions on the data policy and terms of use of the data consisted of a discussion of what the future availability of the data should be. It was agreed to make at least presence/absence data available as soon as a first round of papers based on the combined database was published. Data would be published through EurOBIS (www.marbef.org/data/eurobis), one of the activities of the MarBEF network. From there, the data would be available for inclusion in the International Portal of OBIS (www.iobis.org) and the Global Biodiversity Information Facility (GBIF, www.gbif.org).

Compiling an integrated database on soft-bottom macrobenthic fauna is a time-consuming effort which requires sound and professional data management. However, costs are relatively low, compared to the costs (ship time, sampling, identification) of obtaining the data. A cost-comparison described by Zeller et al. (2005) indicated that recovering paper-based, archived data and then creating electronic databases is much cheaper than the actual survey itself: the price for recovery would only be 0.2 to 0.5% of the original survey costs. As analysis of these integrated data can lead to new scientific insights, data from different sources, obtained for different uses can, in combination, provide substantial added value. An important task when integrating different data is the standardisation of taxonomic, geographical and sampling information. An essential element of this standardisation is the adoption of controlled vocabularies for various aspects of the data, including sampling methodology, gazetteer and taxonomic reference lists. Ideally, these would be used by the scientist while collecting the data. Also, the scientist should be aware

of potential re-use of the data collected; all too often raw data are discarded, or information essential to other uses of the data are not written down. In general, there is a lack of training in, and understanding of, data management issues. The curriculum of environmental sciences (and other data-intensive sciences) should include instruction in proper data management procedures, just like students are now exposed to courses in statistical analysis.

Working with an integrated database such as MacroBen has its pitfalls: combining data from different sources causes extra variability and biases. For such large databases, it might be better to aggregate some of the data, sacrificing precision for accuracy. Since identifications were done by very many different groups, it is impossible to verify whether these identifications were done consistently—this would require ring-testing, or comparable inter-comparison exercises, for which no time or resources were available. It might therefore be beneficial to work at the level of genus or even higher, rather than at the more precise but possibly inaccurate level of species. Nevertheless, the result of the MacroBen data integration was the production of a very large dataset which, with suitable adjustments for inconsistencies in the data, provided enhanced statistical power in the identification of significant patterns. Also, the larger ‘footprint’ of the data (i.e. the larger taxonomic, geographical and temporal cover) should, in practice, result in the presence of stronger signals in the data. This increase in power and in signal should more than compensate for any loss in precision.

Of course, integration of existing data from disparate sources will rarely match the outcome of collaborative international research and monitoring programmes in terms of precision and accuracy. In the present integrated database, distribution of the data is dictated by research needs of the individual studies, and results in strong geographical and temporal bias. Large differences in sample size and replication restrict the type of analysis that can meaningfully be applied to the data. However, integrating all available data seems to be the only way, for the time being, to build databases that are needed to inform environmental management at and above sea-wide scales. Opportunistic re-use of available data can lead to problems, but the papers in this Theme Section prove that useful insights can be gained from them.

The MacroBen database and publications are only a first step in an ongoing process. MarBEF is now integrating datasets on hard-bottom biota, and planktonic datasets; a first round of papers on an integrated database on meiofauna was published (Vandepitte et al. 2009). Eventually, these activities will demonstrate the scientific and societal value of building collaborative

databases. Undoubtedly, better science would be done if all data were made publicly available after the publication of the reports and articles that refer to them. Many journals have a policy that ensures public availability of the involved datasets but in practice it seems to be poorly enforced (with the exception of genetic sequence data, which are routinely made available through GenBank). Scientists who are custodians of relevant data are urged to contact the MarBEF data management team and explore possible collaboration.

Acknowledgements. The MacroBen project was carried out in the framework of the MarBEF Network of Excellence ‘Marine Biodiversity and Ecosystem Functioning’ which is funded by the Sustainable Development, Global Change and Ecosystems Programme of the European Community’s Sixth Framework Programme (contract no. GOCE-CT-2003-505446). This publication is contribution number MPS-09025 of MarBEF. Pieter Deckers (then at VLIZ) and Panayiotis Gotsis (HCMR) contributed to the Visual Basic code of the Access front-end. A. S. Y. Mackie, P. G. Oliver, E. I. S. Rees, A. Koukouras, The Norwegian Oil Industry Association (OLF) and Det Norske Veritas are thanked for additional data contributions.

LITERATURE CITED

- Aleffi F, Bettoso N (2001) Il macrozoobenthos del litorale di Barcola (Golfo di Trieste, Nord Adriatico). *Hydrores Inf* 18: 7–15
- Aleffi F, Orel G, Del Piero D, Vio E (1992) Oxygen condition in the Gulf of Trieste (High Adriatic). In: Vollenweider RA, Marchetti R, Viviani R (eds) *Marine coastal eutrophication*. Elsevier, London, p 431–440
- Aleffi F, Goriup F, Orel G, Zuccarello V (1996) Analysis of macrobenthic community structure in three areas of the Gulf of Trieste. *Ann An Istrske Mediter Stud Hist Nat* 9: 39–44
- Aleffi F, Bettoso N, Solis-Weiss V (2003) Spatial distribution of soft-bottom polychaetes along the western coast of the northern Adriatic Sea (Italy). *Ann Ser Hist Nat* 13:211–222
- Arvanitidis C (2000) Polychaete fauna of the Aegean Sea: inventory and new information. *Bull Mar Sci* 66:73–96
- Arvanitidis C, Koutsoubas D, Dounas C, Eleftheriou A (1999) Annelid fauna of a Mediterranean lagoon (Gialova Lagoon, South-West Greece): community structure in a severely fluctuating environment. *J Mar Biol Assoc UK* 79: 849–856
- Arvanitidis C, Bellan G, Drakopoulos P, Valavanis V, Dounas C, Koukouras A, Eleftheriou A (2002) Seascape biodiversity patterns along the Mediterranean and the Black Sea: lessons from the biogeography of benthic polychaetes. *Mar Ecol Prog Ser* 244:139–152
- Arvanitidis C, Valavanis VD, Eleftheriou A, Costello MJ and others (2006) MedOBIS: biogeographic information system for the eastern Mediterranean and Black Sea. *Mar Ecol Prog Ser* 316:225–230
- Bachelet G (1981) Données préliminaires sur l’organisation trophique d’un peuplement benthique marin. *Vie Milieu* 31:205–213
- Bogdanos C, Simboura N, Zenetos A (2002) The benthic fauna of Geras Gulf, (Lesvos isl. Greece): inventory, distribution and some zoogeographical considerations. *Hell Zool Arch* 6:1–22

- Borja A, Franco J, Pérez V (2000) A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. *Mar Pollut Bull* 40:1100–1114
- Borja A, Muxika I, Franco J (2003) The application of a marine biotic index to different impact sources affecting soft-bottom benthic communities along European coasts. *Mar Pollut Bull* 46:835–845
- Cattijnsse A, Claus S, D'Haenens T, Haspeslagh J, T'Jampens R, Vanden Berghe E, Vandepitte L (2006) IMIS integrated marine information system input manual. Flanders Marine Institute (VLIZ), Oostende, Belgium
- Clarke KR, Warwick RM (1998) A taxonomic distinctness index and its statistical properties. *J Appl Ecol* 35:523–531
- Clarke KR, Warwick RM (1999) The taxonomic distinctness measure of biodiversity: weighting of step lengths between hierarchical levels. *Mar Ecol Prog Ser* 184:21–29
- Clarke KR, Warwick RM (2001) A further biodiversity index applicable to species lists: variation in taxonomic distinctness. *Mar Ecol Prog Ser* 216:265–278
- Cochrane SJ, Dahle S, Oug E, Gulliksen B, Denisenko S (1998) Benthic fauna in the Northern Barents Sea. Akvaplan-niva rep no. 434.97.1286, Norwegian Ministry of Environment, Oslo
- Connor DW, Brazier DP, Dalkin MJ, Hill TO, Holt RHF, Northen KO, Sanderson WG (1999) Marine nature conservation review: marine biotope classification for Britain and Ireland, version 97.06. In: Picton BE, Costello MJ (eds) *The BioMar biotope viewer: a guide to marine habitats, fauna and flora in Britain and Ireland*. Environmental Sciences Unit, Trinity College, Dublin
- Costello MJ, Vanden Berghe E (2006) 'Ocean biodiversity informatics': a new era in marine biology research and management. *Mar Ecol Prog Ser* 316:203–214
- Costello MJ, Emblow CS, Picton BE (1996) Long term trends in the discovery of marine species new to science in Britain and Ireland. *J Mar Biol Assoc UK* 76:255–257
- Costello MJ, Emblow CS, White R (eds) (2001) *European register of marine species. A check-list of the marine species in Europe and a bibliography of guides to their identification*. Patrimoines naturelles, 50 Museum national d'Histoire naturelle, Paris
- Craeymeersch JA, Heip CHR, Buijs J (1997) *Atlas of North Sea benthic infauna: based on the 1986 North Sea Benthos Survey*. ICES Coop Res Rep 218. ICES, Copenhagen
- Cuvellier D, Claus S, Appeltans W, Vanhoorne B, Vanden Berghe E, Costello MJ (2006) European register of marine species (ERMS)—plans turning into reality! *MarBEF Newsl* 4:14–15
- Dahle S, Denisenko S, Denisenko N, Cochrane S (1992) Benthic fauna in the Pechora Sea. *Sarsia* 83:183–210
- Dahle S, Anisimova NA, Palerud R, Pearson TH, Matishov GG (2009) Macrobenthic fauna of the Franz Josef Land archipelago. *Polar Biol* 32:169–180
- Degraer S, Verfaillie E, Willems W, Adriaens E, Vincx M, Van Lancker V (2008) Habitat suitability modelling as a mapping tool for macrobenthic communities: an example from the Belgian part of the North Sea. *Cont Shelf Res* 28:369–379
- Fabri MC, Galéron J, Maudire G (2004) BIOCEAN: a new database for deep-sea benthic ecological data. In: Vanden Berghe E, Appeltans W, Fabricius M (eds) *Abstracts, Ocean Biodiversity Informatics*, Hamburg, Germany, 29 Nov to 1 Dec 2004. Flanders Marine Institute (VLIZ), Oostende, Belgium, p 63
- Fabri MC, Galéron J, Larour M, Maudire G (2006) Combining the Biocean database for deep-sea benthic data with the online Ocean Biogeographic Information System. *Mar Ecol Prog Ser* 316:215–224
- Floen S, Gjøsaeter H, Korneliussen R, Sagen H, Thorvaldsen P, Wennevik V (1993) An integrated database for marine research. ICES CM 1993 (D:23 Sess.T.)
- Grassle JF (2000) The Ocean Biogeographic Information System (OBIS): an on-line, worldwide atlas for accessing, modeling and mapping marine biological data in a multidimensional geographic context. *Oceanography (Wash DC)* 13:5–9
- Gray JS (1981) *The ecology of marine sediments: an introduction to the structure and function of benthic communities*. Cambridge University Press, Cambridge
- Heip C, Basford D, Craeymeersch J, Dewarumez JM and others (1992) The benthic communities of the North Sea: a summary of the results of the North Sea benthos survey. ICES Coop Res Rep 190:148–175
- Herrmann M (2004) Makrozoobenthos — Gemeinschaften arktischer Weichböden: Struktur und Bedeutung als Nahrungsgrundlage demersaler Fische. MSc thesis, Christian-Albrechts-Universität, Kiel
- Hill MO (1979) TWINSPAN. A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of individuals and attributes. Cornell University, Ithaca, NY
- Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 52:577–586
- Hyland J, Balthis L, Karakassis I, Magni P and others (2005) Organic carbon content of sediments as an indicator of stress in the marine benthos. *Mar Ecol Prog Ser* 295:91–103
- Janas U (1998) Influence of hypoxia and hydrogen sulphide on macrofauna in the Gulf of Gdansk. PhD thesis, Gdańsk University
- Kallonas M, Zenetos A, Gofas S (1999) Notes on the ecology and distribution of microbivalvia in Greek waters. *Conchiglia* 291:11–20
- Karakassis I, Eleftheriou A (1997) The continental shelf of Crete: structure of macrobenthic communities. *Mar Ecol Prog Ser* 160:185–196
- Karakassis I, Eleftheriou A (1998) The continental shelf of Crete: the benthic environment. *PSZN I: Mar Ecol* 19:263–277
- Karakassis I, Hatziyanni E (2000) Benthic disturbance due to fish farming analyzed under different levels of taxonomic resolution. *Mar Ecol Prog Ser* 203:247–253
- Karakassis I, Hatziyanni E, Tzapakis M, Plaiti W (1999) Benthic recovery following cessation of fish farming: a series of successes and catastrophes. *Mar Ecol Prog Ser* 184:205–218
- Karakassis I, Tzapakis M, Hatziyanni E, Papadopoulou KN, Plaiti W (2000) Impact of cage farming of fish on the seabed in three Mediterranean coastal areas. *ICES J Mar Sci* 57:1462–1471
- Karakassis I, Tzapakis M, Smith CJ, Rumohr H (2002) Fish farming impacts in the Mediterranean studied through sediment profiling imagery. *Mar Ecol Prog Ser* 227:125–133
- Kendall MA (1995) Are Arctic soft-sediment macrobenthic communities impoverished? *Polar Biol* 16:393–399
- Kendall MA, Widdicombe S (1999) Small scale patterns in the structure of macrofaunal assemblages of shallow soft sediments. *J Exp Mar Biol Ecol* 237:127–140
- Kisseleva MI, Revkov NK, Kopitov YP (1999) The modern state and the long-term changes of zoobenthos in Streletskaya Bay (Sevastopol area). *Hydrobiol J* 35 (original in *Gidrobiologicheskij zhurnal* 1997; 33:3–13 (in Russian))
- Koutsoubas D, Dounas C, Arvanitidis C, Kornilios S, Petihakis G, Triantafyllou G, Eleftheriou A (2000) Macrobenthic community structure and disturbance assessment in a Mediterranean lagoon (Gialova Lagoon, Ionian Sea). *ICES J Mar Sci* 57:1472–1480

- Laudien J, Herrmann M, Arntz WE (2004) Soft bottom community structure and diversity in Arctic Kongsfjord. In: Wiencke C (ed) The coastal ecosystem of Kongsfjorden, Svalbard. Ber Polarforsch 492:103–111
- Mackie ASY, Oliver PG, Rees EIS (1995) BIOMÖR 1 benthic biodiversity in the Southern Irish Sea. BIOMÖR Rep 1, National Museums & Galleries of Wales, Cardiff
- Magurran AE (1998) Ecological diversity and its measurement. Chapman & Hall, London
- McGrath D, Costello MJ, Emblow C (2000) The hermit crab, *Diogenes pugilator* (Roux, 1829) in Irish waters. Biol Environ Proc R Ir Acad B 100:115–118
- Mikhailova TV (1989) Changes by depth in distribution of benthos in Yalta region. In: Long-term changes of zoobenthos in the Black Sea. Naukova Dumka Publ, Kiev, p 18–34 (in Russian)
- Mironov OG, Kiryukhina LN, Divavin IA (1992) Sanitary and biological studies in the Black Sea. S.-Pb. Hydrometeoizdat, St. Petersburg (in Russian)
- Nicolaidou A, Zenetos A, Pancucci MA, Simboura N (1993) Comparing ecological effects of two different types of pollution using multivariate techniques. PSZN I: Mar Ecol 14:113–128
- Occhipinti-Ambrogi A, Savini D, Forni G (2005) Macrobenthos community structural changes off Cesenatico coast (Emilia Romagna, Northern Adriatic), a six-year monitoring programme. Sci Total Environ 353:317–328
- Pancucci MA, Zenetos A (1990) On the echinoderm-fauna of the Geras Gulf. Acta Adriat 31:293–300
- Papathanassiou E, Zenetos A (1993) A case of recovery in benthic communities following a reduction in chemical pollution in a Mediterranean ecosystem. Mar Environ Res 36:131–152
- Papathanassiou E, Bogdanos C, Zenetos A (1989) Application of the log-normal distribution on the benthic communities of Geras Gulf (Mytilene, Aegean Sea). Proc 1st Conf Environmental Science and Technology, p 457–469 (in Greek)
- Petrov AN (1999) Using of bioindication approach for assessment of modern anthro-pogenic impact upon coastal marine environment around Crete (the Aegean Sea) and Crimea (the Black Sea). In: Balopoulos E, et al. (eds) Oceanography of the eastern Mediterranean and the Black Sea, similarities and differences of two interconnected basins, Athens, Greece, 23–26 Feb 1999. Scientific Report, European Commission, p 396–398
- Petrov AN (2000) Responses of the Black Sea macrobenthic communities upon organic enrichment impact of bottom sediments. Ecol Morya 51:45–51 (in Russian with English abstract)
- Petrov AN, Revkov NK (2005) Multivariate analysis of zoobenthos structure under technogenic pollution impact (Streletskaia bay, SW Crimea). Proc Ternopol Pedagogical University, Ser Biol, Spec Issue: Hydrobiology 4: 177–180 (in Russian)
- Philippart CJM (ed) (2007) Impacts of climate change on the European marine and coastal environment: ecosystems approach. ESF Marine Board Position Paper. European Science Foundation, Marine Board, Strasbourg
- Pinedo S, Sarda R, Martin D (1996) Seasonal dynamics and structure of soft-bottom assemblages in the Bay of Blanes (Western Mediterranean Sea). Publ Espec Inst Esp Oceanogr 22:61–70
- Pinedo S, Sarda R, Martin D (1997) Comparative study of the trophic structure of soft-bottom assemblages in the Bay of Blanes (Western Mediterranean Sea). Bull Mar Sci 60: 529–542
- Rabaut M, Guilini K, Van Hoey G, Vincx M, Degraer S (2007) A bio-engineered soft-bottom environment: the impact of *Janice conchilega* on the benthic species-specific densities and community structure. Estuar Coast Shelf Sci 75: 525–536
- Rees HL, Eggleton JD, Rachor E, Vanden Berghe E (2007) Structure and dynamics of the North Sea benthos. ICES Coop Res Rep 288. ICES, Copenhagen
- Revkov NK (2003a) Zoobenthos regional peculiarities—modern condition of biological diversity in near-shore zone of Crimea (the Black Sea sector). Ekosi-Gidrophizika, Sevastopol, p 218–221 (in Russian)
- Revkov NK (2003b) Taxonomical composition of the bottom fauna at the Black Sea Crimean coast—modern condition of biological diversity in near-shore zone of Crimea (the Black Sea sector). Ekosi-Gidrophizika, Sevastopol, p 209–218, 326–338 (in Russian)
- Revkov NK, Nikolaenko TV (2002) Biodiversity of zoobenthos in the coastal zone of the south coast of Crimea (Laspi Bay area). Russ J Mar Biol 28:151–162
- Revkov NK, Sergeeva NG (2004) Current state of the zoobenthos at the Crimean shores of the Black Sea. International Workshop on the Black Sea benthos, 18–23 April 2004, Istanbul, p 189–217
- Rosenberg R, Blomqvist M, Nilsson HC, Cederwall H, Dimming A (2004) Marine quality assessment by use of benthic species-abundance distributions: a proposed new protocol within the European Union Water Framework Directive. Mar Pollut Bull 49:728–739
- Rumohr H, Karakassis I, Nørrevang Jensen J (2001) Estimating species richness, abundance and diversity with 70 macrobenthic replicates from the Western Baltic. Mar Ecol Prog Ser 214:103–110
- Sarda R, Pinedo S, Martin D (1999) Seasonal dynamics of macroinfaunal key species inhabiting shallow soft-bottoms in the Bay of Blanes (NW Mediterranean). Acta Oecol 20: 315–326
- Sarda R, Pinedo S, Grémare A, Taboada S (2000) Changes in the dynamics of shallow soft-bottom due to man-made disturbance processes in the Catalan Western Mediterranean Sea. ICES J Mar Sci 57:1446–1457
- Sergeeva NG (1992) The characteristics of the free-living nematode fauna of Yalta gulf under anthropogenic impact. In: Long-term changes of zoobenthos in the Black Sea. Naukova Dumka, Kiev, p 170–184 (in Russian)
- Sergeeva NG, Mikhailova TV (1989a) Distribution and structure of bottom communities in some regions of the Southern coast of Crimea. VINITI Publ. #1329-B89, Moscow (in Russian)
- Sergeeva NG, Mikhailova TV (1989b) Distribution and structure of bottom communities in Yalta Gulf. VINITI Publ #1328-B89, Moscow (in Russian)
- Sides EM, Picton BE, Costello MJ, Crean E and others (1995) Identification and mapping of marine biotopes. In: Carroll M, Dubsy K (eds) Coastal zone management: from needs to action. Coastwatch Europe, Dublin, p 198–203
- Simboura N, Zenetos A (2002) Benthic indicators to use in ecological quality classification of Mediterranean soft bottoms marine ecosystems, including a new biotic index. Mediterr Mar Sci 3:77–111
- Simboura N, Zenetos A, Panayotidis P, Makra A (1995) Changes in benthic community structure along an environmental pollution gradient. Mar Pollut Bull 30:470–474
- Solis-Weiss V, Rossin P, Aleffi F, Bettoso N, Orel G, Vrišer B (2001) Gulf of Trieste: sensitivity areas using benthos and GIS techniques. In: Ozhan E (ed) Proc 5th Int Conf Mediterranean Coastal Environment, MEDCOAST 01, 23–27 October 2001, Hammamet, Tunisia. Middle East Technical University, Ankara, Turkey, Vol 1–3:1567–1578
- Speybroeck J, Bonte D, Gheskiere T, Grootaert P and others

- (2005) Biologische evaluatie van elf strandzones langs de Vlaamse kust, B.E.ST. Final report. Ministry of the Flemish Community, AMINAL, Nature Department, Brussels
- Triantafyllou G, Petihakis G, Dounas C, Koutsoubas D, Arvanitidis C, Eleftheriou A (2000) Temporal variations on benthic communities and their behaviour to physicochemical forcing. A numerical approach. ICES J Mar Sci 57: 1507–1516
- Van Aartsen JJ, Zenetos A (1987) The genus *Clathromangalia* Monterosato, 1884, in the Mediterranean, with the description of *Clathromangalia ferhi* n.sp. Conchiglia 222–223: 10–11
- Van Hoey G, Degraer S, Vincx M (2004) Macrobenthic communities of soft-bottom sediments at the Belgian continental shelf. Estuar Coast Shelf Sci 59:599–615
- Van Hoey G, Vincx M, Degraer S (2005) Small-scale geographical patterns within the macrobenthic *Abra alba* community. Estuar Coast Shelf Sci 64:751–763
- Van Hoey G, Vincx M, Degraer S (2007) Temporal variability in the *Abra alba* community determined by global and local events. J Sea Res 58:144–155
- Vanden Berghe E (2006) The European register of marine species: tool for data integration. In: Abstracts ICES Annu Sci Conf, 19–23 Sep 2006, Maastricht. ICES, Copenhagen, p 227
- Vanden Berghe E, Rees HL, Eggleton JD (2007) NSBP 2000 data management. In: Rees HL, et al. (ed) Structure and dynamics of the North Sea benthos. ICES Coop Res Rep 288:7–20
- Vandepitte L, Vanaverbeke J, Vanden Berghe E, Bezerra TC (2007) Developing an integrated database to perform joint analyses: MANUELA (Meiobenthic And Nematode biodiversity Unravelling Ecological and Latitudinal Aspects). In: Abstracts 13th Int Meiofauna Conf (THIRIMCO), 29 Jul to 3 Aug 2007, Recife, Brazil. p 17
- Vandepitte L, Vanaverbeke J, Vanhoorne B, Hernandez F, Bezerra TN, Mees J, Vanden Berghe E (2009) The MANUELA database: an integrated database on meiobenthos from European marine waters. Meiofauna Mar 17:35–60
- Vatova A (1949) La fauna bentonica dell'alto e medio Adriatico. Nova Thalassia 1:1–110
- Wasmund N, Pollehne F, Postel L, Siegel H, Zettler ML (2004) Biologische Zustandseinschätzung der Ostsee im Jahre 2003. Meereswiss Ber 60:1–87
- Willems W, Goethals P, Van den Eynde D, Van Hoey G and others (2008) Where is the worm? Predicting the spatial distribution of the tube-building polychaete *Lanice conchilega* (Pallas, 1766). Ecol Modell 212:74–79
- Zeller D, Froese R, Pauly D (2005) On losing and recovering fisheries and marine science data. Mar Policy 29:69–73
- Zenetos A, Papathanassiou E (1989) Community parameters and multivariate analysis as a means of assessing the effects of tannery effluents on the macrobenthos. Mar Pollut Bull 20:176–181
- Zenetos A, Papathanassiou E, Van Aartsen JJ (1991) Analysis of benthic communities in the Cyclades plateau (Aegean Sea) using ecological and paleoecological datasets. PSZN I: Mar Ecol 12:123–137
- Zenetos A, Bogdanos C, Simbhora N (1992) Comparative study of the macrozoobenthos in the Gulfs of Geras and Kalloni. In: Proc 3rd Pan-Hellenic Symp Oceanography and Fisheries, Athens. p 548–558
- Zenetos A, Simbhora N, Papathanassiou E (1993) Biological indices in a volcanic area in S. Aegean (Santorini). In: Proc 4th Pan-Hellenic Symp Oceanography and Fisheries, Rhodes. p 226–229
- Zenetos A, Simbhora N, Panayotidis P (1994) Effects of sewage on the distribution of benthic fauna in Saronikos Gulf. UNEP MAP Tech Rep Ser 80:39–72
- Zenetos A, Christianidis S, Pancucci MA, Simbhora N, Tziavos CH (1997) Oceanologic study of an open coastal area in the Ionian Sea with emphasis on its benthic fauna and some zoogeographical remarks. Oceanol Acta 20: 437–451
- Zenetos A, Revkov NK, Konsulova T, Sergeeva N, Simbhora N, Todorova V, Zaika VE (1999) Coastal benthic diversity in the Black and Aegean Seas. Mediterr Mar Sci 1/2: 105–117

Appendix 1. Datasets included in the MarBEF database

Aleffi I.F., 2004: LBMRev dataset. Marine Biology Laboratory, University of Trieste, Italy.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=602

The primary aim for this data collection was to build a database and to apply GIS techniques in the Gulf of Trieste and North Adriatic Sea (Aleffi et al. 1992, 1996, 2003, Aleffi & Bettoso 2001, Solis-Weiss et al. 2001). *Dataset code: lm.*

Arvanitidis C., 2000: Aegean polychaetes dataset. Hellenic Centre for Marine Research, Greece.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=595

The Hellenic Centre for Marine Research collected samples in the context of a number of National and EU-funded projects. The datasets provide qualitative temporal information on polychaete worms in the Aegean and Ionian Sea (Arvanitidis 2000, Arvanitidis et al. 2002, 2006). *Dataset code: ap.*

Bachelet G., 2004: Benthos Gironde Estuary. Université Bordeaux 1, Station Marine d'Arcachon, Laboratoire d'Océanographie Biologique, France.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=597

The data are from 3 sources, namely the impact of dredging on benthic macrofauna (1979 data), a comparison of macrobenthos between major European tidal estuaries (1991–1992 data) and ecological monitoring of macrobenthos in an area affected by the effluents of a nuclear power plant (2004 data) (Bachelet 1981). *Dataset code: ba.*

Craeymeersch J.A., Kingston P., Rachor E., Duineveld G., Heip C.H.R., Vanden Berghe E., 1986: North Sea Benthos Survey.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=67

The data are the result of the 1986 North Sea Benthos Survey, an activity of the Benthos Ecology Working Group of ICES. Benthic samples were taken in a standardised way, on a regular grid covering the whole of the North Sea, and analysed by scientists from 10 laboratories. Extensive work was done to standardise taxonomy and identifications across the different laboratories (Heip et al. 1992, Craeymeersch et al. 1997). *Dataset code: ns.*

Dahle S., Cochrane S.J., 1992: Northern Barents Sea 1992.

Appendix 1. (continued)

Akvaplan-niva, Norway.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=12

Benthic fauna in the Northern Barents Sea was mapped in July 1992 by Akvaplan-niva (Cochrane et al. 1998). *Dataset code: o2.*

Dahle S., Palerud R., Anisimova N., 1992: Benthic fauna around Franz-Josef Land. Akvaplan-niva, Norway.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=11

Benthic fauna in the Franz-Josef Land was mapped in August 1992 by Akvaplan-niva (Dahle et al. 2009). *Dataset code: o8.*

Dahle S., Cochrane S.J., Denisenko S., 1992: Benthic fauna around Pechora Sea. Akvaplan-niva, Norway.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=461

Benthic fauna in the Pechora Sea was mapped in July 1992 by Akvaplan-niva (Dahle et al. 1992). *Dataset code: o7.*

Degraer S., Van Hoey G., Willems W., Speybroeck J., Vincx M., 2003: MacroDat Belgium. Macrobenthic data from the Belgian part of the North Sea from 1976 onwards. Ghent University, Biology Department, Marine Biology Section, Belgium.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=633

All data on macrobenthos of the Belgian part of the North Sea collected within several projects of the University of Ghent are integrated into this dataset. It was originally compiled to assure a proper data management of all physically separated datasets on macrobenthos (Van Hoey et al. 2004, 2005, 2007, Speybroeck et al. 2005, Rabaut et al. 2007, Van Hoey et al. 2007, Degraer et al. 2008, Willems et al. 2008). *Dataset code: ug.*

Dounas C., 1988: Benthos Cretan Continental Shelf. Hellenic Centre for Marine Research, Greece.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=598

This pilot benthic study was carried out during the summer of 1988 on the continental Cretan shelf (Koutsoubas et al. 1990, Karakassis & Eleftheriou 1997, 1998). *Dataset code: do.*

Dounas C., 1996: Environmental impact assessment of oil pollution accident in Gialova lagoon and Navarino Bay. Hellenic Centre for Marine Research, Greece.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=272

This dataset was part of a national project on the consequences of the Navarino oil spill in 1994 at the Gialova Lagoon, Ionian Sea (Arvanitidis et al. 1999, Koutsoubas et al. 2000, Triantafyllou et al. 2000). *Dataset code: M0.*

Fabri M.C., 2001: Biocean dataset. IFREMER France.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=490

This dataset combines information on deep sea benthic fauna collected during several IFREMER cruises from 1967 to 2001. Due to its large geographical scope and the deep sea habitat, this dataset can be considered to some extent as an outlier of the MacroBen database (Fabri et al. 2004, 2006). *Dataset code: bo.*

Gray J.S., 1997: Oslofjord 96-97. University of Oslo, Norway. Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=1148

Unpublished soft bottom benthic data collected in the Oslofjord within the framework of a PhD study. Dataset code: of.

Grémare A., Amouroux J.M., Labruno C., 1998: Redit dataset. Observatoire Océanologique de Banyuls-Sur-Mer, Laboratoire d'Océanographie Biologique, France.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=213

This dataset describes the benthic macrofauna in the Gulf of Lyons. It was collected in order to assess the spatial scale of the increase of the polychaete *Ditropa arietina*. *Dataset code: gr.*

Herrmann M., Laudien J., 2004: Soft bottom community structure and diversity in Arctic Kongsfjorden. Alfred-Wegener-Institute for Polar and Marine Research, Bremerhaven, Germany.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=614

The data were collected as part of a diver-operated shallow-water (5-30m) research study of the soft bottom benthic community in Arctic Kongsfjord/Spitsbergen (Laudien et al. 2004, Herrmann 2004). *Dataset code: ko.*

Hummel H., Sistermans W.C.H., Escaravage V., 2004: BIS dataset of the south-western part of Netherlands (1985-2004). Benthos information System, Monitor Taskforce - NIOO-CEME, Netherlands.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=599

This was the largest contributing dataset (about 130 000 distribution records), over 95 % of which was sourced from several Dutch monitoring programmes carried out in the Dutch Delta area (Eastern- & Western Scheldt Estuary, Lake Grevelingen, Lake Veere) and the Dutch coastal zone (Voordelta). *Dataset code: ni.*

Janas U., 2002: Gulf of Gdansk dataset. University of Gdansk, Institute of Oceanography, Department of Marine Biology and Ecology, Poland.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=612

The datasets relate to 2 studies in the Gulf of Gdansk, one in 1995 to study the influence of hypoxia and hydrogen sulphide on the macrofauna and the other in 2002 to analyse the distribution of the macrofauna (Janas 1998). *Dataset code: gd.*

Karakassis I., 1997: Cretan shelf macrofauna and macrofauna of Greek polluted sites. Biology Department, University of Crete, Greece.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=613

The original objective of the Cretan Shelf dataset was to unravel the structure of the benthic ecosystem of the continental shelf of Crete; the macrofauna of Greek polluted sites gathered data on the benthic effects of fish farming (Karakassis & Eleftheriou 1997, 1998, Karakassis et al. 1999, 2000, 2002, Karakassis & Hatzizanni 2000, Hyland et al. 2005). *Dataset code: ka.*

Kendall M.A., 1996: Arctic soft-sediment macrobenthos. Plymouth Marine Laboratory, UK.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=533

Samples of the macrobenthic fauna of soft sediments were collected from around Svalbard during the 1991 Arctic EPOS

Appendix 1. (continued)

cruise of RV Polarstern (Kendall 1995). *Dataset code: ar.*

Kendall M.A., Widdicombe S., 1999: Plymouth Sound dataset. Soft sediment macrobenthos from the Plymouth Sound from 1995. Plymouth Marine Laboratory, UK. Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=534

The intended use of this dataset of the Plymouth Sound was the description of small-scale spatial patterns in soft sediment fauna (Kendall & Widdicombe 1999). *Dataset code: pl.*

Kotwicki L., 1996: Bay of Puck dataset. Polish Academy of Sciences, Institute of Oceanology, Department of Marine Ecology, Poland.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=611

The dataset was collected in July 1996 for habitat mapping of the Bay of Puck. *Dataset code: pu.*

Koukouras A., 2000: Northern Aegean dataset. Aristotelian University of Thessaloniki Department of Zoology and Zoological Museum, School of Biology, Greece.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=630

The dataset consists of a museum collection of long term qualitative macrobenthic data from the Aegean and Ionian Sea. An extensive list of peer-reviewed articles was derived from this dataset. *Dataset code: o1.*

Mackie A.S.Y., Oliver P.G., Rees E.I.S., 1991: Biomôr 1 dataset. Benthic data from the Southern Irish Sea from 1989–1991. National Museum and galleries of Wales, Cardiff, UK. Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=1600

Surveys of benthic invertebrates of the Southern Irish Sea were carried out in 1989 and 1991. Both quantitative and qualitative samples were taken for faunal and sediment analysis (Mackie et al. 1995). *Dataset code: o5.*

Nevrova H., Petrov A., Revkov N., 1996: Laspibay-Black Sea dataset—IBSS, Sevastopol. Institute of Biology of the Southern Seas (IBSS), Ukraine.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=629

The dataset arose from an investigation of the structure of macrozoobenthos and microphytobenthos assemblages in relation to recreation and marifarming development in Laspi Bay (SW Crimea, the Black Sea) (Petrov 2000, Revkov & Nikolaenko 2002, Revkov 2003a,b, Revkov & Sergeeva 2004). *Dataset code: m9.*

Occhipinti-Ambrogi A., 2002: Cesenatico dataset. Department of Ecologia del Territorio, University of Pavia, Italy.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=600

This dataset was compiled in order to monitor long-term responses of the macrobenthos community to environmental quality changes in the Northern Adriatic Sea (Occhipinti-Ambrogi et al. 2005). *Dataset code: oc.*

Olsfjord F., 1994: Oslofjord 93-94. Akvaplan-niva, Norway. Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=1152

Unpublished data collected as part of a general environmental investigation of the Oslofjord. *Dataset code: of.*

Petrov A., Revkov N., 1986: MegFeod-Black Sea dataset—IBSS, Sevastopol. Institute of Biology of the Southern Seas

(IBSS), Ukraine.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=626

The dataset arose from a comparative investigation of the structure and spatial patterns of macrozoobenthic assemblages along the coast of southeastern Crimea (Sergeeva & Mikhailova 1989a, Mironov et al. 1992, Petrov 1999, 2000, Hyland et al. 2005). *Dataset code: m1.*

Petrov A., Revkov N., 1993: Strelbay-Black Sea dataset—IBSS, Sevastopol. Institute of Biology of the Southern Seas (IBSS), Ukraine.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=627

The dataset arose from an investigation of the structure and condition of macrozoobenthos assemblages under persistent anthropogenic impact (Kisseleva et al. 1999, Petrov 2000, Revkov 2003a,b, Petrov & Revkov 2003, 2005). *Dataset code: m4.*

Petrov A., Revkov N., 1986: Jalta-Black Sea dataset—IBSS, Sevastopol. Institute of Biology of the Southern Seas (IBSS), Ukraine.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=628

The dataset arose from comparative studies of the condition of macro- and meio-zoobenthos around Jalta Gulf under the influence of domestic sewage discharge and recreation activity (Sergeeva 1992, Mikhailova 1989, Petrov 1999, 2000). *Dataset code: m5.*

Picton B.E., Emblow C.S., Morrow C.C., Sides E.M., Tierney P., McGrath D., McGeough G., McCrea M., Dinneen P., Falvey J., Dempsey S., Dowse J., Costello M.J., 1999: Marine sites, habitats and species data collected during the BioMar survey of Ireland. Environmental Sciences Unit, Trinity College, Dublin, Ireland.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=345

Through marine habitat and biotope surveys, the Biomar project (1992–1996) aimed at improved management of coastal ecosystems in Britain and Northern Ireland. Qualitative benthic data from 200 littoral and 700 sublittoral sites around the Republic of Ireland were collected both by direct observations on seashores and by SCUBA divers (Sides et al. 1995, Costello et al. 1996, Connor et al. 1999, McGrath et al. 2000). *Dataset code: bi.*

Rumohr H., 1995: Kiel Bay intercalibration dataset. Leibniz Institute of Marine Sciences, Marine Ecology Division, Germany.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=618

The ICES/HELCOM Intercalibration Exercise provided a series of replicate samples from 1 station in Kiel Bay from May 1995 (Rumohr et al. 2001). *Dataset code: 70.*

Rumohr H., Fleischer D., 2004: N3 data of Kiel Bay. Leibniz Institute of Marine Sciences, Marine Ecology Division, Germany.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=617

Long term monitoring of all invertebrate species of station N3 in Kiel Bay in the Western Baltic was conducted between 1986 and 2004. *Dataset code: n3.*

Sardá R., 1997: Soft Bottom Communities of the Bay of

Appendix 1. (continued)

Blanes. Centre d'Estudis Avançats de Blanes (CEAB) Consejo Superior de Investigaciones Científicas (CSIC), Spain.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=603

Soft bottom communities of the Bay of Blanes in Girona were sampled from 1992 to 1997 in the framework of 2 Spanish national projects analysing the regulation of biological production in the Mediterranean littoral (Pinedo et al. 1996, 1997, Sarda et al. 1999, 2000). *Dataset code: bl*.

The Norwegian Oil Industry Association, 2001: Offshore reference stations, North/Norwegian sea. The Norwegian Oil Industry Association (OLF), Akvaplan-niva and Det Norske Veritas, Norway.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=999

The dataset is a compilation of selected reference stations from large-scale annual surveys of the benthic fauna in the vicinity of petroleum installations, funded by the Norwegian Oil Industry Association. *Dataset code: o3*.

The Norwegian Oil Industry Association, 2002: Offshore reference stations, Norwegian/Barents Sea. The Norwegian Oil Industry Association (OLF), Akvaplan-niva and Det Norske Veritas, Norway.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=997

Compilation of selected reference stations from large-scale annual surveys of the benthic fauna in the vicinity of petroleum installations, funded by the Norwegian Oil Industry Association. *Dataset code: o4*.

The Norwegian Oil Industry Association, 2000: Offshore reference stations, Finnmark. The Norwegian Oil Industry Association (OLF), Akvaplan-niva and Det Norske Veritas, Norway.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=998

Compilation of selected reference stations from large-scale annual surveys of the benthic fauna in the vicinity of petroleum installations, funded by the Norwegian Oil Industry Association. *Dataset code: o6*.

Vatova A., 1936: Fauna Bontonica dataset, Instituto di Biologia Marine per l'Adriatico, Venezia, Italy.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=596

These data were collected between 1934 and 1936 as part of a PhD study. Scientists from the Hellenic Centre for Marine Research computed the dataset afterwards (Vatova 1949). *Dataset code: fb*.

Węśławski J.M., Malec E., Jaskuła R., Włodarska-Kowalczyk M., Kędra M., 2002: Polish Arctic Marine Programme. Macro-benthic data from Hornsund from 2002. Polish Academy of Sciences, Institute of Oceanology, Poland.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=243

This dataset was collected as part of an All Taxa Biodiversity Inventory (ATBI). *Dataset code: hs*.

Whomersley P., 2003: National Marine Monitoring Programme. Benthos data of the North Sea, Irish Sea, English Channel from 2002–2003. Centre for environment, fisheries and aquaculture science, Burnham-on-Crouch, UK.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=615

The data were generated as part of the UK's mandatory monitoring requirements under the OSPAR Joint Assessment and Monitoring Programme (JAMP) and to monitor long-term trends at selected temporal stations around the UK. *Dataset code: np*.

Zenetos A., 1986: Zoobenthos Kyklades, Aegean Sea dataset. Hellenic Centre for Marine Research, Greece.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=625

Under the framework of the project 'Oceanographic studies in open Seas', a survey of the benthic fauna of the Central Aegean Sea was performed in July 1986 (Van Aartsen & Zenetos 1987, Zenetos et al. 1991, 1993, Kallonas et al. 1999). *Dataset code: m8*.

Zenetos A., 1987: Aegean Sea coastal benthic communities, Geras Gulf (Mytilini), chemical pollution (tannery effluents). Hellenic Centre for Marine Research, Greece.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=622

An impact study of a tannery effluent on the benthic community was conducted over 2 and a half years in the Geras Gulf (Papathanassiou et al. 1989, Zenetos & Papathanassiou 1989, Pancucci & Zenetos 1990, Zenetos et al. 1992, Papathanassiou & Zenetos 1993, Bogdanos et al. 2002). *Dataset code: m2*.

Zenetos A., 1990: Seasonal zoobenthos, Saronikos. Hellenic Centre for Marine Research, Greece.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=289

This dataset was collected under the framework of 2 different projects, recording biological parameters and an investigation of the effects of a wastewater treatment plant on the zoobenthos in Saronikos Gulf. (Nicolaidou et al. 1993, Zenetos et al. 1994, 1999, Simboursa et al. 1995, Simboursa & Zenetos 2002). *Dataset code: m3*.

Zenetos A., 1991: Kalamitsi dataset. Hellenic Centre for Marine Research, Greece.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=623

The dataset arose from a survey of the Kalamitsi area of the Ionian Sea (Zenetos et al. 1997). *Dataset code: m6*.

Zenetos A., Bogdanos C., 1992: Kerkyra dataset. Hellenic Centre for Marine Research, Greece.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=624

The dataset arose from a survey of the Kerkyra area of the Ionian Sea. *Dataset code: m7*.

Zettler M.L., 2005: Macrozoobenthos Baltic Sea (1980–2005) as part of the IOW-Monitoring. Institut für Ostseeforschung Warnemünde, Germany.

Available at: www.marbef.org/modules.php?name=People&lvl=Data&show=html&dasid=601

This time-series study of species composition and biomass or abundance of macrozoobenthos in the Belt Sea and the Baltic proper provide data from 1979 to 2005 as part of the HELCOM monitoring programme (Wasmund et al. 2004). *Dataset code: ze*.