Mobilizing historical marine data within EMODnet Biology

Vasilis Gerovasileiou[1], Stamatina Nikolopoulou[1], Dimitra Mavraki[1], Christos Arvanitidis[1], Leen Vandepitte[2], Gizem Poffyn[2], Joana Beja[2], Ruben Perez Perez[2], Laura Boicenco[3], Marina Lipizer[4], Menashè Eliezer[4]

[1] HCMR - Greece, [2] VLIZ - Belgium, [3] NIMRD - Romania, [4] OGS - Italy

EMODnet BIOLOGY *Dive into data on Europe's marine life*

RATIONALE

Historical marine data are often scattered in old publications, reports and expedition logbooks, either in the form of hard copies or in simple and unorganized spreadsheets in electronic storage media.

Such sources provide **invaluable** biological and environmental data that could be used to reconstruct and model past conditions or predict trends and shifts in distribution range, biological invasions or even regional species extinctions (Faulwetter *et al*. 2016).

This is extremely important for European Seas and adjacent marine regions, which are particularly vulnerable to an increasing number of human activities and pressures.

Recognizing the importance of rescuing and mobilizing historical marine data, **EMODnet Biology** has developed a *long-term data archaeology and rescue strategy*, under a dedicated work package (WP3).

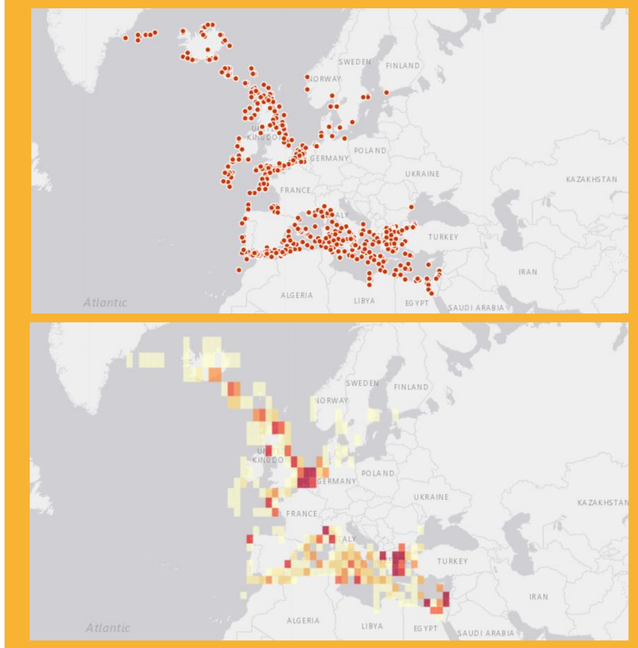
Overall objective:

To fill spatial and temporal gaps in aquatic species occurrences and make the rescued historical data available freely through the EMODnet portal and global biogeographic and biodiversity information systems (e.g., OBIS and its regional nodes, WoRMS and its sub-registers, LifeWatch Species Information Backbone).

Special focus:

Understudied, "data poor" regions which are susceptible to environmental alterations and biological invasions such as the South-Eastern Mediterranean and Black Seas.

DATA RESCUE



Distribution of historical occurrence data published in OBIS. Colour shades represent the number of records (higher values in red).

Identification & prioritization of data sources

- ✓ **240 marine historical** (1890s to 1950s) and **rescue datasets** (1960s to 2000s) were assembled by HCMR, focusing mainly on the **Mediterranean Sea** and adjacent marine areas.
- ✓ **95 datasets** were identified for rescue by VLIZ, ranging from 1930s to mid-1990s, describing data collected mainly in the **North Sea** area.
- ✓ **Paper archives and primary protocols** from the **Black Sea** waters were scanned and rescued by NIMRD, containing data on phytoplankton, zooplankton and macrozoobenthos dating back to the 1950s, which remained unavailable to the wider research community.

*A set of **purpose-build criteria** was adopted for the prioritization and selection of datasets to be rescued and mobilized, including thematic and taxonomic cover, temporal and geographic scope, language and readiness in their availability (e.g., online digital files versus hard copies in libraries).*

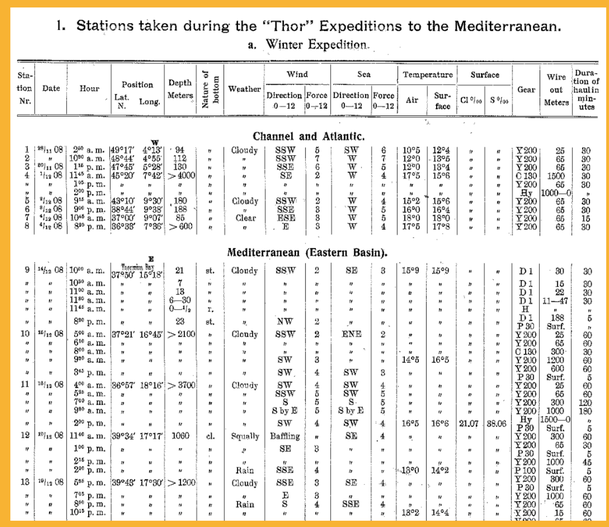
Data rescue & challenges

So far **130 historical marine datasets** (1868-1999) have been rescued and published, which included **139,740 occurrence records** of **5,099 taxa**.

Several challenges were identified at all stages of the data "life-cycle", from dataset identification and metadata extraction to the digitization and quality control of historical data (Mavraki *et al*. 2016), such as:

- ✓ lack of standardization,
- ✓ lack of georeferencing accuracy,
- ✓ taxonomic inconsistencies and updates,
- ✓ misspellings of taxa and locations,
- ✓ poorly documented or missing sampling protocols.

Facing these challenges is of paramount importance since loss of such data equals to loss of unique resources required to understand global changes.



I. Stations taken during the "Thor" Expeditions to the Mediterranean.

a. Winter Expedition.

Station Nr.	Date	Hour	Position Lat. N.	Position Long.	Depth Meters	Nature of bottom	Weather	Wind Direction 0—12	Wind Force 0—12	Sea Direction 0—12	Sea Force 0—12	Temperature Air	Temperature Surface	Surface Cl ‰	Surface S ‰	Gear	Wire out Meters	Duration of haulin minutes
Channel and Atlantic.																		
1	30/11 08	2^00 a. m.	49°17′	4°13′	94	"	Cloudy	SSW	5	SW	6	10°5	12°4	"	"	Y 200	25	30
2	"	10^00 a. m.	48°44′	4°55	112	"	"	SSW	7	W	7	12°0	13°6	"	"	Y 200	65	30
3	30/11 08	1^00 p. m.	47°45′	6°28′	120	"	"	SSE	4	W	6	12°0	13°4	"	"	Y 200	65	30
4	1/12 08	11^45 a. m.	45°20′	7°42′	> 4000	"	"	SE	2	W	4	17°5	15°6	"	"	C 130	1500	30
"	"	1^00 p. m.	"	"	"	"	"	"	"	"	"	"	"	"	"	Y 200	65	30
"	"	2^00 p. m.	"	"	"	"	"	"	"	"	"	"	"	"	"	Hy	1000—0	"
5	2/12 08	9^45 a. m.	43°10′	9°30′	180	"	Cloudy	SSW	2	W	4	15°2	16°6	"	"	Y 200	65	30
6	3/12 08	9^00 p. m.	38°41′	9°35′	188	"	"	SSE	3	W	5	16°0	16°4	"	"	Y 200	65	30
7	4/12 08	10^45 a. m.	37°00′	9°07′	85	"	Clear	ESE	3	W	5	18°0	18°0	"	"	Y 200	65	15
8	4/12 08	8^00 p. m.	36°38′	7°36′	> 600	"	"	E	3	W	4	17°5	17°8	"	"	Y 200	65	30
Mediterranean (Eastern Basin).																		
9	14/12 08	10^00 a. m.	Kastelli Bay 37°50′	16°18′	21	st.	Cloudy	SSW	2	SE	3	15°9	15°9	"	"	D 1	30	30
"	"	10^00 a. m.	"	"	"	"	"	"	"	"	"	"	"	"	"	D 1	16	30
"	"	11^30 a. m.	"	"	13	"	"	"	"	"	"	"	"	"	"	D 1	22	30
"	"	11^45 a. m.	"	"	6—30	"	"	"	"	"	"	"	"	"	"	D 1	11—47	30
"	"	11^45 a. m.	"	"	0—14	T.	"	"	"	"	"	"	"	"	"	H	"	"
"	"	8^00 p. m.	"	"	23	st.	"	NW	2	"	"	"	"	"	"	D 1	188	5
10	16/12 08	5^00 a. m.	37°21′	16°45′	> 2100	"	Cloudy	SSW	2	ENE	2	"	"	"	"	P 30	Surf.	"
"	"	6^45 a. m.	"	"	"	"	"	"	"	"	"	"	"	"	"	Y 200	25	60
"	"	8^00 a. m.	"	"	"	"	"	"	"	"	"	"	"	"	"	Y 200	65	60
"	"	9^00 a. m.	"	"	"	"	"	SW	3	"	"	16°5	16°5	"	"	C 130	300	30
"	"	3^45 p. m.	"	"	"	"	"	SW	4	SW	3	"	"	"	"	Y 200	1200	60
11	18/12 08	4^00 a. m.	36°57′	18°16′	> 3700	"	Cloudy	SW	4	SW	4	"	"	"	"	Y 200	600	60
"	"	5^00 a. m.	"	"	"	"	"	SSW	5	SW	5	"	"	"	"	P 30	Surf.	5
"	"	7^45 a. m.	"	"	"	"	"	S	5	S	5	"	"	"	"	Y 200	25	60
"	"	9^00 a. m.	"	"	"	"	"	S by E	5	S by E	5	"	"	"	"	Y 200	65	60
"	"	2^00 p. m.	"	"	"	"	"	SW	4	SW	4	16°5	16°6	21,07	38,06	Y 200	300	120
12	19/12 08	11^00 a. m.	39°34′	17°17′	1080	cl.	Squally	Baffling	"	SE	4	"	"	"	"	Y 200	1000	180
"	"	1^00 p. m.	"	"	"	"	"	SE	3	"	"	"	"	"	"	Hy	1500—0	"
"	"	2^15 p. m.	"	"	"	"	"	"	"	"	"	"	"	"	"	P 30	Surf.	5
"	"	2^45 p. m.	"	"	"	"	Rain	SSE	4	"	"	15°0	14°2	"	"	P 100	Surf.	5
13	20/12 08	5^45 p. m.	39°45′	17°30′	> 1200	"	Cloudy	SSE	3	SE	4	"	"	"	"	Y 200	300	60
"	"	7^45 p. m.	"	"	"	"	"	E	3	"	"	"	"	"	"	P 30	Surf.	5
"	"	8^00 p. m.	"	"	"	"	Rain	S	4	SSE	4	"	"	"	"	Y 200	1000	60
"	"	10^15 p. m.	"	"	"	"	"	"	"	"	"	18°2	14°4	"	"	Y 200	65	60
"	"	"	"	"	"	"	"	"	"	"	"	"	"	"	"	Y 200	15	60

Original presentation of sampling metadata of the historical "Thor" Expedition to the Mediterranean Sea and adjacent waters (from: Schmidt 1912 Høst & Son, Copenhagen, 49 pp.).

NEXT STEPS

In **Phase IV** of EMODnet Biology (2021-2023):

- ✓ An **inventory of historical marine datasets** will be created to form the basis of a targeted digitization effort.
- ✓ When identified datasets concern paper files, they will be transferred to a **digital format** before the data content is transformed to a DarwinCore format.
- ✓ A set of historical biological datasets will be used to evaluate the feasibility of **engaging volunteers** to data rescue.
- ✓ **Recognition** of specific **ecological traits** and sampling devices/methodologies in the text of historical documents will be investigated.

References

Faulwetter et al. (2016) EMODnet Workshop on mechanisms and guidelines to mobilise historical data into biogeographic databases. *Research Ideas and Outcomes* 2: e10445.

Mavraki et al. (2016) Rescuing biogeographic legacy data: The "Thor" Expedition, a historical oceanographic expedition to the Mediterranean Sea. *Biodiversity Data Journal* 4: e11054.

Acknowledgments

This work has been financially supported by the EC DG-MARE (EMODnet Observation and Data network - Lot5 – Biology) and by the Greek Government under the GSRT and the National Strategic Reference Framework (LifeWatchGreece Research Infrastructure (ESFRI) project - MIS 384676). The Belgian archaeology work was carried out in the framework of LifeWatch, of which the Flemish contribution is funded through FWO (Research Foundation Flanders).

Online resources:

- Course on How to contribute datasets to EMODnet Biology (https://classroom.oceanteacher.org/course/view.php?id=430)

bio@emodnet.eu

@EMODnet
@EurOBIS_VLIZ
@lifewatchgreece
@LifeWatchVLIZ

The European Marine Observation and Data Network (EMODnet) is financed by the European Union under Regulation (EU) No 508/2014 of the European Parliament and of the Council of 15 May 2014 on the European Maritime and Fisheries Fund