European Commission

JRC TECHNICAL REPORTS

Transitional waters North East Atlantic geographic intercalibration group

Benthic invertebrate fauna ecological assessment methods

Muxica I, Van Hoey G, Bonne W, Salas Herrero F

2019

Joint Research Centre

EUR 29644 EN

Contact information
Name: Fuensanta Salas Herrero
Address: Via Enrico Fermi 2749, 21027 Ispra (VA), Italy
Email: Fuensanta.Salas-Herrero@ec.europa.eu
Tel.: +39 0332 78 5701

JRC Science Hub

https://ec.europa.eu/jrc

JRC115479

EUR 29644 EN

PDF ISBN 978-92-79-99321-3 ISSN 1831-9424 doi:10.2760/784291

Luxembourg: Publications Office of the European Union, 2019

How to cite: Muxica I, Van Hoey G, Bonne W, Salas Herrero F, Transitional waters North East Atlantic Geographic Intercalibration Group. Benthic invertebrate fauna ecological assessment methods; EUR 29644EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-79-99321-3, doi:10.2760/784291, JRC115479

Table of contents

Foreword 1
Acknowledgements 2
Authors 2
Contributors 2
Abstract 3
1. Introduction 4
2. Available data 5
3. Description of national assessment methods 5
3.1 Methods and required BQE parameters 5
3.2 Sampling and data processing 8
3.3 National reference conditions 10
3.4 National boundary setting 12
3.5 Results of WFD compliance checking 12
4. Results IC feasibility checking 14
4.1 Typology 14
4.1.1 Optimizing typology 14
4.2 Pressures addressed 29
4.2.1 Available pressure-response information 29
4.2.2 Pressure-response relations within the common dataset 39
4.2.3 Conclusions 53
4.3 Assessment concept 53
5. Collection of IC dataset and benchmarking 54
5.1 Dataset description 54
5.2 Common benchmark: IC reference conditions or alternative benchmark 56
5.3 Benchmark standardisation 56
6. Comparison of methods and boundaries 57
6.1 IC option and common metrics 57
6.2 Results of the regression comparison 57
6.3 Comparability criteria 58
7. Final results to be included in the EC Decision 60
8. Correspondence common types versus national types 60
9.Gaps of the current intercalibration 60
10.Ecological characteristics 61
Description of reference or alternative benchmark communities 61
Description of good status communities 61
11. Conclusions 62
References 63

List of abbreviations and definitions 65
Key Terms: 65
Abbreviations: 65
List of figures... 67
List of tables... 70
ANNEX I... 72

Foreword

The European Water Framework Directive (WFD) requires the national classifications of good ecological status to be harmonised through an intercalibration exercise. In this exercise, significant differences in status classification among Member States are harmonized by comparing and, if necessary, adjusting the good status boundaries of the national assessment methods.

Intercalibration is performed for rivers, lakes, coastal and transitional waters, focusing on selected types of water bodies (intercalibration types), anthropogenic pressures and Biological Quality Elements. Intercalibration exercises are carried out in Geographical Intercalibration Groups - larger geographical units including Member States with similar water body types - and followed the procedure described in the WFD Common Implementation Strategy Guidance document on the intercalibration process (European Commission, 2011).

The Technical report on the Water Framework Directive intercalibration describes in detail how the intercalibration exercise has been carried out for the water categories and biological quality elements. The Technical report is organized in volumes according to the water category (rivers, lakes, coastal and transitional waters), Biological Quality Element and Geographical Intercalibration group. This volume addresses the intercalibration of the Transitional Waters-North East Atlantic Benthic Invertebrates Fauna ecological assessment methods.

Acknowledgements

This work for the WFD intercalibration obligation of the Member States has received funding within the framework of JPI Oceans through the JPI Oceans pilot action "Joint funding of the Scientific Intercalibration exercise for the EU Water Framework Directive (WFD) coastal and transitional waters in the North-East Atlantic".

This exercise has been supported by JRC and a review panel formed by recognized experts on IC procedure and benthic invertebrates element.

Authors

Muxica, I., Van Hoey, G., Bonne, W., Salas Herrero, F.

Contributors

Borja, A., Philips, G., Miles, Al., Dubois, S., Desroy, N., Buchet, R., Ximenes, MC., O'Beirn, F., Witt, J., Heyer, K., van Loon, W., Ruiter, H., Neto, J.,Marques, JC., Garcia, E., Puente, A

Abstract

This report gives a technical description on the intercalibration (IC) process of the different benthic assessment approaches for soft sediment habitats (muds to sands) in transitional waters in the North East Atlantic Geographical Intercalibration Group (NEA-GIG) for type NEA 11 (Transitional Waters). Eight member states are involved: Belgium (BE), France (FR), Germany (DE), Ireland (RoI), the Netherlands (NL), Portugal (PT), Spain (SP), and United Kingdom (UK). In Spain, the competent authorities for the WFD application are the regions, as such, for the benthic macroinvertebrates assessment methods three regions have been considered: Andalusia (SP-An), Basque Country (SP-BC) and Cantabria (SP-C). Those member states proposed 7 approaches for IC: AeTV (DE), BAT (PT), BEQI (BE), BEQI2 (NL), IQI (RoI and UK), M-AMBI (DE and SP-BC), QSB (SP-C) and TAsBeM (SP-An). However, AeTV and BEQI are not intercalibrated as they assess benthic invertebrates at water body and ecosystem level, respectively, whereas the rest of methods assess the benthic status at sample level.

1. Introduction

This report gives a technical description on the intercalibration (IC) of the different benthic assessment approaches for soft sediment habitats (muds to sands) in transitional waters in the North East Atlantic Geographical Intercalibration Group (NEA-GIG) for type NEA 11 (Transitional Waters). Eight member states are involved: Belgium (BE), France (FR), Germany (DE), Ireland (RoI), the Netherlands (NL), Portugal (PT), Spain (SP), and United Kingdom (UK). In Spain, the competent authorities for the WFD application are the regions, as such, for the benthic macroinvertebrates assessment methods three regions have been considered: Andalusia (SP-An), Basque Country (SP-BC) and Cantabria (SP-C).

In previous phases of IC a huge effort was done on compiling a database including biological data and standardizing such data, EQR-values and pressure information. Such information was used to define six transitional water body types and some unsuccessful IC exercises were carried out on them.

As some of the results in the end of the 2nd phase of the IC looked promising, a JPI oceans pilot action (http://www.jpi-oceans.eu/intercalibration-eu-water-framework-directive) was called in order to make prospection for progress in the IC of tools for the assessment of benthic invertebrates from transitional waters. The objectives of such an action were:

- WFD method compliance documentation checking, explanations on the justifications for assessing methods including specific parameters, reference conditions and boundary setting procedure. Check or improve pressure-response relationships (2nd phase results are available).
- Check and improve pressure-response relationships for benchmarking (2nd phase result is available), cleaning of dataset and relationships necessary.
- Improve benchmark standardization for regional differences.
- Improve comparability analyses.
- Prepare progress intercalibration technical report.

This report compiles all the information regarding the benthic assessment approaches, boundary- and reference settings for each Member State and common dataset characteristics. Specific analyses were conducted to demonstrate the pressure-response relationships of the benthic assessment approaches, detect possible bio-geographical differences in the common dataset, perform a continuos benchmarking standarization and the comparability analyses following the intercalibration guidelines (Guidance document 14: guidance document on the intercalibration process 2008-2011).

2. Available data

For this exercise two databases were used: (1) NEAGIG TW benthos intercalibration-v9-2011-05-05 DRAFT and (2) NEAGIG_MS_EQR-pressures08september2011-forIC.

The former is the last available version of the database prepared by the national experts involved in the 2nd phase of IC. It includes abundance and biomass data at species level standardized by sampling area and sieving mesh size, for homogenization.

The latter is the last available version of the EQR calculations and also includes pressure information.

NEAGIG TW benthos intercalibration-v9-2011-05-05 DRAFT database includes 9,635 "samples". Some of such "samples" are replicates from a single sampling site and survey. It also includes samples from Sweden, although they no longer declare transitional waters. This database was cleaned up by removing Swedish data and combining (average) replicated "samples". This led to a cleaned up database containing 6,665 samples (included as a separate file: Abundances_cleared_database.xlsx).

NEAGIG_MS_EQR-pressures08september2011-forIC database includes 6,668 samples, also including Swedish data. This database was cleaned up by removing such samples. This led to a database containing 6,580 samples (included as a separate file: EQR-pressures_cleared_database.xlsx).

3. Description of national assessment methods

3.1 Methods and required BQE parameters

Within the NEA-GIG region for transitional waters there were 8 different benthic assessment tools defined. However, as it will be explained below, not all of them were considered for this IC exercise. Most member states participating in the IC proposed a method, except France which has currently no final benthic assessment tool for transitional waters (NEA11). United Kingdom and Ireland are the only member states sharing a method (IQI).

Table 1 Overview of the national assessment methods

Member State	Method	Included in this IC exercise?
Belgium	Benthic Ecosystem Quality Index (BEQI)	No
Germany (DE)	Aestuar Type Verfahren (AeTV)	No
Germany (DE)	Multivariate AZTI's Marine Biotic Index (M-AMBI)	Yes
Netherlands (NL)	Benthic Ecosystem Quality Index-2 (BEQI2)	Yes
Portugal (PT)	Benthic Assessment Tool (BAT)	Yes
Spain - Andalusia (SP-An)	Taxonomically Sufficient Benthic Multimetric (TaSBeM)	No
Spain - Basque Country (SP-BC)	Multivariate AZTI's Marine Biotic Index (M-AMBI)	Yes
Spain - Cantabria (SP-C)	Quality of Soft Bottoms (QSB)	Yes
Ireland & United Kingdom (RoI & UK)	Infaunal Quality Index (IQI)	No

BEQI: multilevel approach (Ecosystem, Habitat, Community), which aims integrating the functional role of benthos and the vulnerability to physical changes in the environment. At ecosystem level BEQI uses the relationship between macrobenthic biomass and system productivity; at habitat level, BEQI considers the spatial distribution of habitats within an ecosystem and addresses the diversity of habitat types; at community level BEQI evaluates the benthic macrofauna community per habitat or ecotope, based on the number os species, total density, total biomass and similarity (Bray-Curtis). This method assesses the quality of benthic invertebrate at ecosystem level, whereas the remainder of methods (but AeTV) give an EQR at sample level.

AeTV: assessment of the oligohaline stretches of the transitional waters based on the presence/absence of indicator species, with mean species number and Fisher's α-diversity as co-metrics. This method assesses the quality of benthic invertebrate at water body level, but the remainder of methods (but BEQI) give an EQR at sample level.

BEQI2: multimetric index calculated including AMBI, Shannon index (H') and species richness (S) as follows (Van Loon *et al.*, 2015):

$$\mathrm{BEQI2} = \frac{1}{3} \times \left[\frac{(\mathrm{S_{ass}} - \mathrm{S_{bad}})}{(\mathrm{S_{ref}} - \mathrm{S_{bad}})}\right] + \frac{1}{3} \times \left[\frac{(\mathrm{H'_{ass}} - \mathrm{H'_{bad}})}{(\mathrm{H'_{ref}} - \mathrm{H'_{bad}})}\right] + \frac{1}{3} \times \left[\frac{(\mathrm{AMBI_{ass}} - \mathrm{AMBI_{bad}})}{(\mathrm{AMBI_{ref}} - \mathrm{AMBI_{bad}})}\right],$$

where *ass*, *bad* and *ref* indicate the respective values in the assessed sample, for bad conditions and for reference conditions.

BAT: calculated from a factor analysis including AMBI, Margalef index and Shannon index (Teixeira *et al.*, 2009).

TaSBeM: multimetric index calculated using species BO2A index (Dauvin and Ruellet, 2009) and Margalef index at family level as follows:

$$\mathrm{TaSBeM} = 0.56 \times \frac{\log 2 - \mathrm{BO2A_{obs}}}{\log 2 - \mathrm{BO2A_{ref}}} + 0.44 \times \frac{\mathrm{dfam_{obs}}}{\mathrm{dfam_{ref}}},$$

where *dfam* equals the Margalef index calculated at family level, and *obs* and *ref* indicate the observed and reference values for each of the metrics.

M-AMBI: calculated from a factor analysis including AMBI, Shannon index and species richness (Muxika *et al.*, 2007). Germany applied the M-AMBI for the status of benthic invertebrates at least at mesohaline and polyhaline stretches, calculated using their own boundaries.

QSB: multimetric index calculated using species richness (S), community composition (Bc) and structure (Bs), abundance of opportunistic species (Op), excess of total abundance (N^+) and deficit of total abundance (N^-) as follows:

$$\mathrm{QSB} = \frac{\mathrm{S_{EQR}} + \mathrm{Bcs_{EQR}} + \mathrm{Op_{EQR}}}{3} \times \mathrm{N^{+}}_{\mathrm{EQR}} \times \mathrm{N^{-}}_{\mathrm{EQR}},$$

where S_{EQR} equals the species richness divided by the reference value, Bcs_{EQR} equals the sum of the benthic composition (Bray-Curtis distance between assessing sample and a reference community, using presence/absence data) divided by the reference value and the benthic structure (Bray-Curtis distance between assessing sample and a reference community, using abundance data) divided by the reference value, Op_{EQR} equals the reference value divided by the relative abundance of opportunistics, N^{+}_{EQR} equals the reference value divided by the total abundance and N^{-}_{EQR} equals the total abundance divided by the reference value (Puente *et al.*, 2010).

IQI: multimetric index calculated including AMBI, Simpson's diversity index (1-λ') and species richness (S) as follows:

$$\mathrm{IQI} = \left(0.38 \times \left(\frac{1 - (\mathrm{AMBI}/7)}{1 - (\mathrm{AMBI_{ref}}/7)}\right) + 0.38 \times \left(\frac{1 - \lambda'}{1 - \lambda'_{\mathrm{ref}}}\right) + 0.54 \times \left(\frac{\mathrm{S}}{\mathrm{S_{ref}}}\right)^{0.1} - 0.4\right) / 0.6,$$

where *ref* indicates the reference values for each of the indices (Phillips *et al.*, 2014).

In table 2, a brief overview of the metrics included in each of the national assessment methods is shown.

Table 2 Overview of the metrics included in the national assessment methods

Member State	Full BQE method	Composition	Abundance	Disturbance sensitive taxa	Diversity	Biomass	Taxa indicative of pollution	Combination rule of metrics
DE	M-AMBI	Not strictly – only as groups of different sensitivity, but also including richness of all species	Not strictly – only relative abundance of different sensitivity groups (in AMBI)	Included in AMBI	Included	Not included	Included in AMBI	Factor Analysis for an ordination of samples after AMBI, species richness and Shannon index values, and vectorial distance to reference conditions
NL	BEQI2	Not strictly – only as groups of different sensitivity, but also including richness of all species	Not strictly – only relative abundance of different sensitivity groups (in AMBI)	Included in AMBI	Included	Not included	Included in AMBI	See previous section
PT	BAT	Not strictly – only as groups of different sensitivity, but also including richness of all species	As Margalef index and also relative abundances of different sensitivity groups (in AMBI)	Included in AMBI	Included	Not included	Included in AMBI	Factor Analysis for an ordination of samples after AMBI, Margalef index and Shannon index values, and vectorial distance to reference conditions
SP-An	TaSBeM	Not strictly – only as groups of different sensitivity (in BO2A)	As Margalef index and also relative abundances of different sensitivity groups (in BO2A)	Included in BO2A (amphipods, which are considered sensitive, except genus *Jassa*)	Margalef index at family level	Not included	Included in BO2A (opportunistic annelids)	See previous section
SP-BC	M-AMBI	Not strictly – only as groups of different sensitivity, but also including richness of all species	Not strictly – only relative abundance of different sensitivity groups (in AMBI)	Included in AMBI	Included	Not included	Included in AMBI	Factor Analysis for an ordination of samples after AMBI, species richness and Shannon index values, and vectorial distance to reference conditions
SP-C	QSB	As Bray-Curtis similarity to reference composition	Included	Percentage of opportunistic species (EG IV and V from AMBI)	Included as species richness	Not included	Included	See previous section
RoI & UK	IQI	Not strictly – only as groups of different sensitivity, but also including richness of all species	Not strictly – only relative abundance of different sensitivity groups (in AMBI	Included in AMBI	Included	Not included	Included in AMBI	See previous section

3.2 Sampling and data processing

In table 3, a brief overview of the sampling and data processing for each of the national assessment methods is shown. Wider information on this issue was compiled during WISER project and is available online (http://www.wiser.eu/results/method-database/).

Table 3 Overview of the sampling and data processing of the national assessment methods included in the IC exercise

	DE	NL	PT	SP-An	SP-BC	SP-C	RoI & UK
- Sampling/survey device	Intertidal: tube Subtidal: Van Veen	Box corer	Van Veen	Grab	Intertidal: spade Subtidal: Van Veen	Intertidal: corer Subtidal: Van Veen	Intertidal: corer Subtidal: grab
- How many sampling/survey occasions (in time) are required to allow for ecological quality classification of survey site or area?	1 per year (some 2 per year)	1 per year	1 per year	1 per year	1 per year	1 per year	1 for classification
- Sampling/survey months	Spring and late summer/autumn	Autumn	Spring	Any	Winter	Summer	Late Winter/Early Spring
- Which method is used to select the sampling /survey site or area?	Representativeness within water body, based on expert knowledge	Random/Stratified	Expert knowledge	Expert knowledge	Representativeness within water body (stratified by salinity), based on expert judgment	Expert knowledge/ Stratified	Random/Stratified
- How many spatial replicates per sampling/ survey occasion are required to allow for ecological quality classification of sampling/ survey site or area?	Intertidal: 10 replicates Subtidal: 10 replicates	Several, until 0.21-2.97 m^2 (habitat-type dependent) are sampled over a period of 3 years	3-5 replicates	3 replicates	3 replicates	Intertidal: 2 replicates Subtidal: 6 replicates	Variable according to habitat, number of years/ stations, methodology and required confidence
- Total sampled area or volume, or total surveyed area, or total sampling duration on which ecological quality classification of sampling/survey site or area is based	Intertidal: 0.2 m^2 Subtidal: 1 m^2	0.1 m^2 (using data pooling)	>0.2 m^2	0.025/0.0675/0.1125 m^2	Intertidal: 0.75 m^2 Subtidal: 0.3 m^2	Intertidal: 0.5 m^2 Subtidal: 0.102 m^2	Intertidal: 0.01 m^2 Subtidal: 0.1 m^2
- Short description of field sampling/survey procedure and processing (sub-sampling)	Samples sieved (1 mm mesh size, mud 0.5 mm), fixed in formalin, and identified using stereo microscope	Samples sieved (1 mm mesh size), fixed in formalin and benthic organisms identified using loupe/microscope	Samples sieved (1 mm mesh size), fixed in formalin and benthic organisms identified using loupe/microscope	Samples sieved (0.5 mm mesh size), fixed in formalin and benthic organisms identified	Samples sieved (1 mm mesh size), fixed in formalin and benthic organisms identified	Samples sieved (1 mm mesh size), fixed in formalin and benthic organisms identified	Samples sieved (0.5 mm mesh size), fixed in formalin and benthic organisms identified

3.3 National reference conditions

In European estuaries there are no real undisturbed sites and historical data are not easily accessible (Borja *et al.*, 2004). Therefore, Member States used different approaches to derive reference conditions (expert knowledge, historical data, modelling). Such methods have been described in table 4 and the values for reference conditions are summarized in table 5.

Table 4 Overview of the methodologies used to derive the reference conditions for the national assessment methods included in the IC exercise

Member State	Type and period of reference or alternative benchmark conditions	Number of reference or benchmark sites	Location of reference/ benchmark sites	Reference criteria used for selection of reference or benchmark sites
DE	Expert knowledge, historical data, least disturbed conditions	No sites, just a reference species list	No sites	No sites
NL	Historical data (1990-2005), modeling	No sites	No sites	No sites
PT	Expert knowledge, historical and monitoring data, and modelling	No sites, just reference conditions	No sites	No sites
SP-An	Alternative benchmark conditions derived by averaging the values of the TaSBeM components at alternative benchmark sites	40T0060, 40T1020 (not provided for IC), 62T0050 and 51T0120 (identified as Bonanza in the IC database)	(ETRS89) 40T0060: -7,4066 ºE / 37,1891 ºN 40T1020: -7,324 ºE / 37,2081 ºN 62T0050: -7,0685 ºE / 37,2103 ºN 51T0120: -6,3431 ºE / 36,8002 ºN	Least disturbed conditions, historical data and expert knowledge; pressure criteria: t-LUSI<4 Soluble P in sediment < 10 mg/kg
SP-BC	Expert knowledge, historical data and modelling (Muxika *et al.*, 2007)	No sites, just reference conditions for several biological communities	No sites	No sites
SP-C	Historical data from least disturbed conditions	No sites, just reference conditions	No sites	Least disturbed sites, without sewage discharges
RoI & UK	Expert knowledge, historical data, least disturbed conditions, modelling	No sites	No sites	No sites

In the 2nd phase of IC, the approach followed by UK for setting their reference conditions was tested as a method to standardise reference conditions for the remainder of the methods. Such an approach consisted in regression models calculated for the common metrics within the different MS methods (Shannon index, species richness, AMBI, etc.) using the quantitative particle size analysis and salinity information as predictor variables, and the metric of interest as the response. However, due to the limited number of data, among other reasons, this approach was not finally agreed as a suitable method to be adopted for standardising reference conditions.

Table 5 Reference conditions defined for each of the methods and types (see table 9)

Method	Type	Habitat/Ecotope/Community	Reference Conditions										
			N^-	N^+	AMBI	BO2A	Bc	Bs	d	Op	S	H'	1-λ'
BAT	**D/F**	Euhaline	-	-	0.8	-	-	-	5.0	-	-	4.1	-
		Polyhaline – Sand	-	-	1.0	-	-	-	4.0	-	-	4.0	-
		Polyhaline – Sandy mud	-	-	1.5	-	-	-	4.0	-	-	4.0	-
		Polyhaline – Muddy sand	-	-	2.4	-	-	-	3.0	-	-	3.8	-
BEQI2	**D**	Mesohaline – Intertidal	-	-	0.57	-	-	-	-	-	29	3.3	-
		Mesohaline – Subtidal	-	-	0.54	-	-	-	-	-	22	3.2	-
		Polyhaline – Intertidal	-	-	1.20	-	-	-	-	-	41	3.6	-
		Polyhaline - Subtidal	-	-	0.63	-	-	-	-	-	31	3.8	-
IQI	**D/E/F**	Euhaline Subtidal – Sand/Mud	-	-	0.28	-	-	-	-	-	78.6	-	1.02
M-AMBI (DE)	**D/E**	Mesohaline – Intertidal – Sand	-	-	0.36	-	-	-	-	-	17	2.98	-
		Mesohaline – Intertidal – Mud	-	-	1.65	-	-	-	-	-	11	2.83	-
		Mesohaline – Subtidal - shallow <6 m	-	-	0.11	-	-	-	-	-	17	2.95	-
		Mesohaline – Subtidal – deep >6 m	-	-	0.03	-	-	-	-	-	14	3.26	-
		Polyhaline – Intertidal – Sand	-	-	1.65	-	-	-	-	-	20	2.86	-
		Polyhaline – Intertidal – Mud	-	-	3.1	-	-	-	-	-	11	1.8	-
		Polyhaline – Subtidal – deep + shallow	-	-	0.07	-	-	-	-	-	34	3.71	-
M-AMBI (SP)	**D/E/F**	*Cerastoderma edule-Scrobicularia plana*	-	-	2.8	-	-	-	-	-	13	2.5	-
		Venus fasciata	-	-	2.0	-	-	-	-	-	32	3.8	-
		Abra alba	-	-	2.1	-	-	-	-	-	40	3.5	-
		Pontocrates arenarius-Eurydice pulchra	-	-	1.0	-	-	-	-	-	9	2.0	-
QSB	**E**	*Abra alba*	297	1127	-	-	80	80	-	10	30	-	-
		Scrobicularia plana	84	481	-	-	80	80	-	10	11	-	-
		Abra tenuis	34	578	-	-	80	80	-	10	15	-	-
TAsBeM	**D/E/F**	Type D – Polyhaline – Mud	-	-	-	0.02	-	-	1.78[a]	-	-	-	-
		Type D – Euhaline – Mud	-	-	-	0.02	-	-	1.78[a]	-	-	-	-
		Type D – Euhaline – Sand	-	-	-	0.00	-	-	2.17[a]	-	-	-	-

[a]Margalef index derived at family level.

* N-/N+: deficit/excess of total abundance; Bc/Bs: community composition/structure; d: Margalef index; Op: abundance of opportunistic species; S: number of species; H': Shannon index; 1-λ': Simpson's diversity index.

3.4 National boundary setting

Table 6 Explanations for national boundary setting of the national methods included in the Ic exercise

Member State	Type of boundary setting: Expert judgment – statistical – ecological discontinuity – or mixed for different boundaries?	Specific approach for HG boundary	Specific approach for GM boundary	BSP: method tested against pressure
DE	Expert judgement	Calibrated against pre-classified sampling sites	Calibrated against pre-classified sampling sites	Tested against pressures
NL	Equidistant division of EQR gradient	Equidistant division of EQR gradient	Equidistant division of EQR gradient	Tested against pressures
PT	Taken from coastal phase I IC	Taken from coastal phase I IC	Taken from coastal phase I IC	Tested against pressures
SP-An	Statistical	25th percentile of TaSBeM values for alternative benchmark samples	25th percentile of TaSBeM values for all samples with t-LUSI ≤ 3(excluding benchmark samples)	Tested against pressures
SP-BC	Taken from coastal phase I IC	Taken from coastal phase I IC	Taken from coastal phase I IC	Tested against pressures
SP-C	Equidistant division of EQR gradient	Equidistant division of EQR gradient	Equidistant division of EQR gradient	Tested against pressures
RoI & UK	Taken from coastal phase I IC	Taken from coastal phase I IC	Taken from coastal phase I IC	Tested against pressures

3.5 Results of WFD compliance checking

The conclusions on compliance checking for the methods included in the IC have been summarized in Table 7**Error! Reference source not found.**. All methods could be considered compliant.

Table 7 List of the WFD compliance criteria and the WFd compliance checking process and results of the national methods included in the IC exercisse

Compliance criteria	Compliance checking conclusions
1. Ecological status is classified by one of **five classes** (high, good, moderate, poor and bad).	Yes
2. High, good and moderate ecological status are set in line with the WFD's **normative definitions** (**Boundary setting procedure**)	Yes
3. **All relevant parameters** indicative of the biological quality element are covered (see Table 1 in the IC Guidance)?	Yes. But: taxonomic composition is used indirectly (except in QSB, for which reference species lists were defined); abundance is used indirectly to calculate different metrics (QSB is the only method with reference value for abundance).

4. Assessment is adapted to **intercalibration common types** that are defined in line with the typological requirements of the Annex II WFD and approved by WG ECOSTAT?	Yes
5. The water body is assessed against **type-specific near-natural reference conditions?**	No. Alternative benchmark conditions (based on a "least disturbed condition" criteria) had to be defined due to the absence of near-natural reference conditions in the intercalibrated type.
6. Assessment results are expressed as **EQRs?**	Yes
7. Sampling procedure allows for **representative** information about water body quality/ecological status **in space and time?**	In most cases, the monitoring is considered as representative by the Member State itself (see annex 1). This aspect is not confirmed by specific, standardized analyses to test their representativeness. Sampling procedures are outlined in general, but not linked with the running WFD monitoring programs.
8. All data relevant for assessing the biological **parameters** specified in the WFD's normative definitions are covered by the **sampling procedure?**	Yes, for all benthic assessment approaches. The sampling procedure defined by each Member State allows the collection of species-abundance data, which is necessary to calculate all metrics of the different benthic assessment approaches.
9. Selected taxonomic level achieves adequate **confidence and precision** in classification?	Yes, for all benthic assessment approaches, with some difference in taxonomic detail per Member State, but sufficient comparability. Taxonomy between Member States datasets is standardized for intercalibration purposes.

4. Results IC feasibility checking

4.1 Typology

4.1.1 Optimizing typology

In the 2nd phase of IC, 6 transitional water body types were defined for NEA. However, the high *intra-type* heterogeneity in the data prevented from obtaining successful results in subsequent analyses. Therefore, in order to reduce the dataset to homogeneous groups of samples (in terms of benthic assemblages) and improve the pressure-response relationships multivariate analyses were carried out on fourth root-transformed benthic abundance data. This task was carried out using the Abundances_cleared_database.xlsx.

For that two families of multivariate analyses were carried out:

- Metric scaling techniques:
    - Two way indicator species analysis (TWInSpAn): in summary, an ordination of samples is first carried out based on their species composition, and species are then grouped based on whether they are in the same samples. Finally, both classifications are combined in a table which can be represented as a dendrogram with the samples grouped after the obtained classification and the indicator species which presence or absence defines each of the groups. This analysis was carried out using WinTWINS 2.3 software (Hill y Šmilauer, 2005).
    - Correspondence Analysis (CA): this analysis allows to plot in an n-dimensional space the ordination of samples and/or species. In order to reduce the effect of rare species, the "downweighting of rare species" option was activated. The scaling was focused on inter-sample distances and Hill's scaling was used. This analysis was carried out using Canoco for Windows 4.5 software (Hill, 1973; Hill y Gauch, 1980).
- Non-metric scaling techniques: for this exercise, Bray-Curtis similarity based on fourth root transformed abundances was used for these analyses. Due to the large size of the database, problems related to computational power were faced. These problems were solved by using the computing grid available in AZTI and by programming some of the commands in R (the scripts are included as separate files: pvclust_bcdist.R and R_scripst.docx).
    - Cluster analysis: in summary, samples are grouped in clusters based on the similarity between them, so that samples with similar species composition tend to be together in the same group. Bootstrap techniques were applied in order to look for "robust" groups.
    - Non-metric multidimensional scaling (nMDS): in summary, samples are projected in an n-dimensional space, based on the similarity among them.

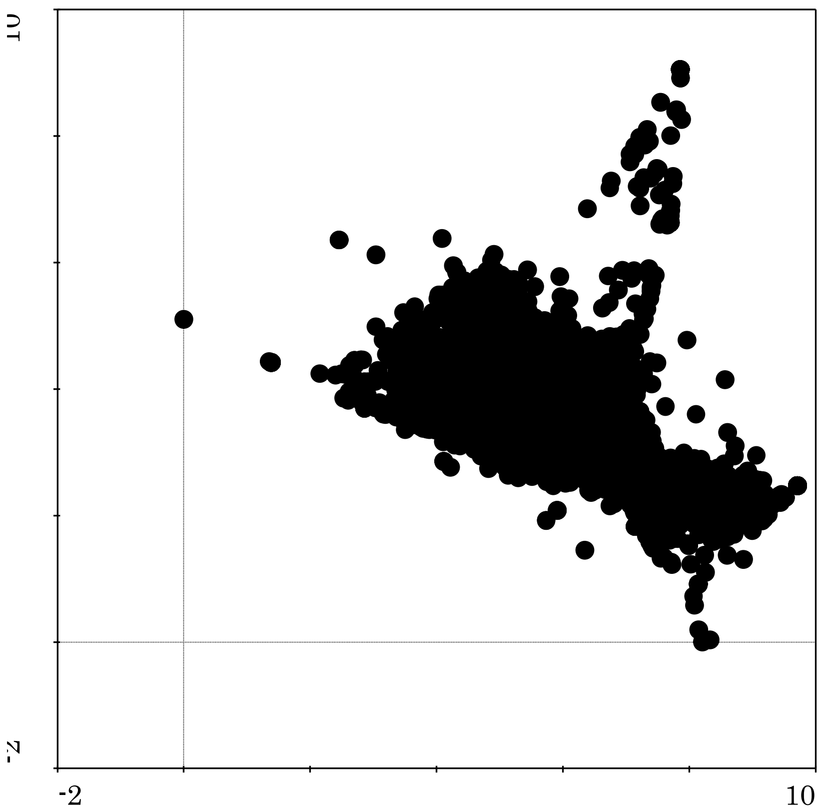
Regarding metric scaling techniques, and after removing outlier samples, a TWInSpAn was carried out. In this analysis four sample groups were distinguished (the resulting groups are included as a separate file: Twinspan_2.pun).

One of the groups included only some French data and a sample from the Netherlands. The other three groups (Final_Twinspan.xlsx) were quite big heterogeneous groups, including samples coming from different countries, estuaries, salinity stretches, sediment types, etc.

Regarding CA, and after removing outlier samples, a detrended correspondence analysis (DCA) was carried out in order to check the length of gradient and decide whether a CA should be carried out (length of gradient>4) or whether a principal component analysis would be an adequate approach (length of gradient<3).

The length of the gradient was 9.7 for the first axis and the ordination plot obtained from the analysis did not show any individual sample falling far from the bulk or outlier.

Figure 1 ordination plot obtained from the detrended correspondece analysis carried out once outliers were removed.



Being the length of gradient>4, as abovementioned, a CA was carried out.

The results show that the first four axes explained less than 12% of variance of species data.

Table 8 Results obtained from the correspondence analysis carried out once outliers were removed

Axes	1	2	3	4	Total inertia
Eigenvalues	0.676	0.557	0.515	0.458	18.871
Cumulative percentage variance of species data	3.6	6.5	9.3	11.7	

As metric-scaling techniques were not able to find homogeneous groups of samples in terms of species composition, a cluster analysis was carried out.

For this, Bray-Curtis distances were calculated using fourth root transformed abundances. As abovementioned, bootstrap techniques were applied in order to check which groups were robust, selecting those with an approximately unbiased p-value higher than 95% (1,000 iterations). This led to 892 groups of 2 to 54 samples, which was considered an unmanageable number of groups.

When the acceptable level for robust groups was set at p=0.75, the size (number of samples) of some groups was increased, but the number of clusters also increased, leading to a total of 1,226 clusters with 2 to 91 samples. Clusters including 50 or more samples were taken (to reduce the number of groups) included samples taken from few water bodies and, most of the times, from single Member States (these clusters are included as a separate file: Clusters.xlsx).

When samples were plotted in an nMDS after the Bray-Curtis distances calculated, they occupied all available 3-dimensional space, not showing any clustering (Figure 2). Moreover, the minimum stress value for the ordination was high (0.333), indicating that the points were close to being arbitrarily placed in the 3-dimensional ordination (Clarke and Warwick, 2001).

As abovementioned, six main types of transitional waters were defined in the 2nd phase of IC, on the basis of their main features (Table 9). Among such transitional water types, only types D, E and F were considered for IC.

When the samples in the nMDS above were identified according to the water body type, groups were not found (Figure 3).

This exercise was repeated identifying the samples according to the saline stretch in which they were located (Figure 4), and according to the tidal level (intertidal/subtidal) (Figure 5).

Figure 2. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.333.

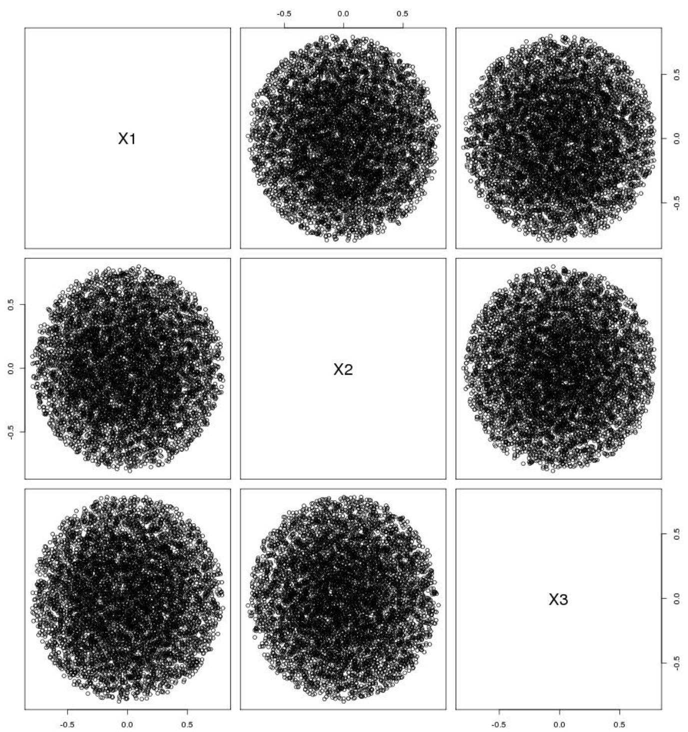


Table 9. Main common IC types identified in NEA transitional water bodies.

Common IC type	Type characteristics	MS sharing IC common type
A	Lagoons	SP, and RoI & UK
B	Freshwater-oligohaline, medium river flow	SP
C	Mesotidal estuary with irregular river flow	PT and SP
D	Large estuaries	DE, NL, PT, SP, and RoI & UK
E	Small-medium estuary with >50% intertidal area	DE, SP, and RoI& UK
F	Small-medium estuary with <50% intertidal area	PT, SP, and RoI & UK

Figure 3. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.333. Black: samples from Type D water bodies; Red: samples from Type E water bodies; Green: samples from Type F water bodies.

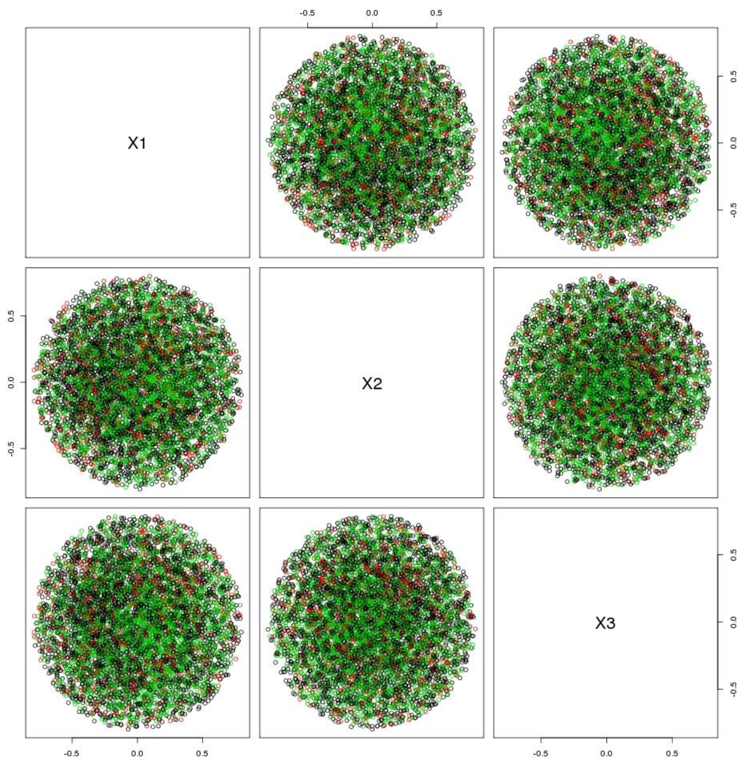


Figure 4. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.333. Black: samples from fresh water stretches; Red: samples from oligohaline stretches; Green: samples from mesohaline stretches; Blue: samples from polyhaline stretches; Cyan: samples from euhaline water stretches.

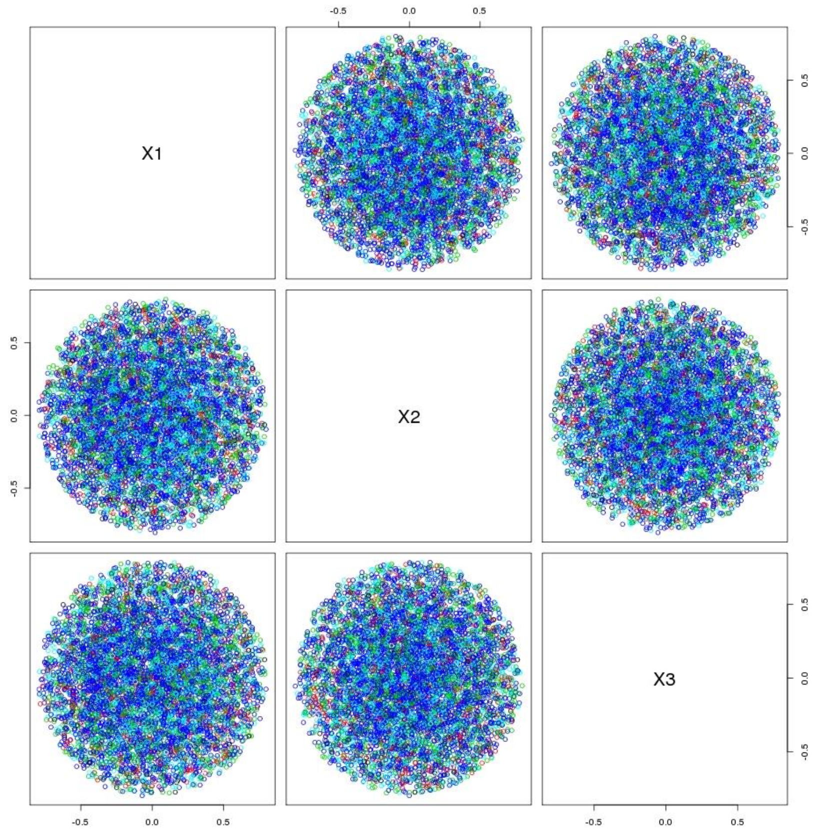
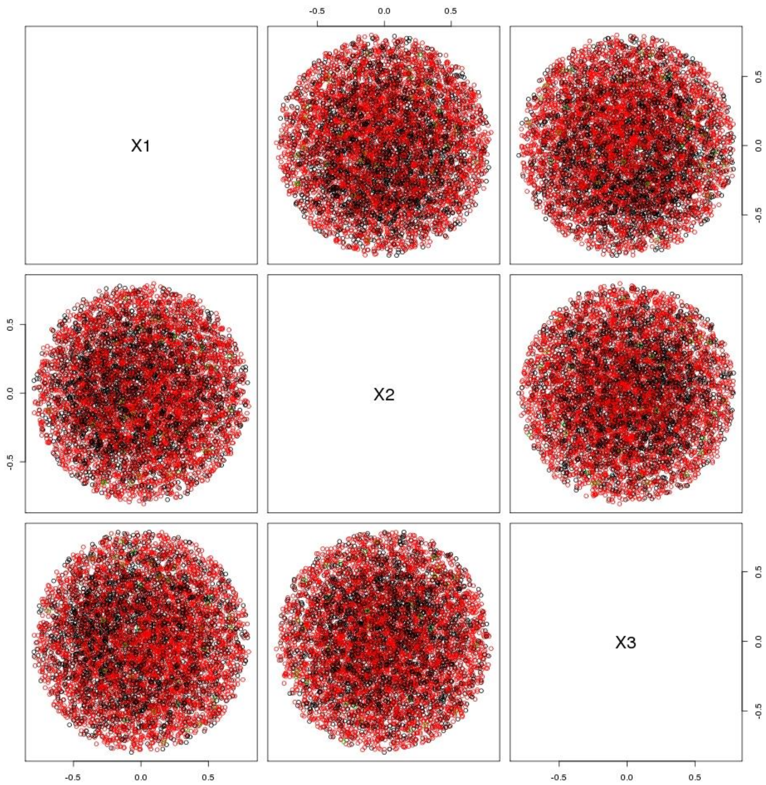


Figure 5. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.333. Black: samples from intertidal sampling stations; Red: samples from subtidal sampling stations.



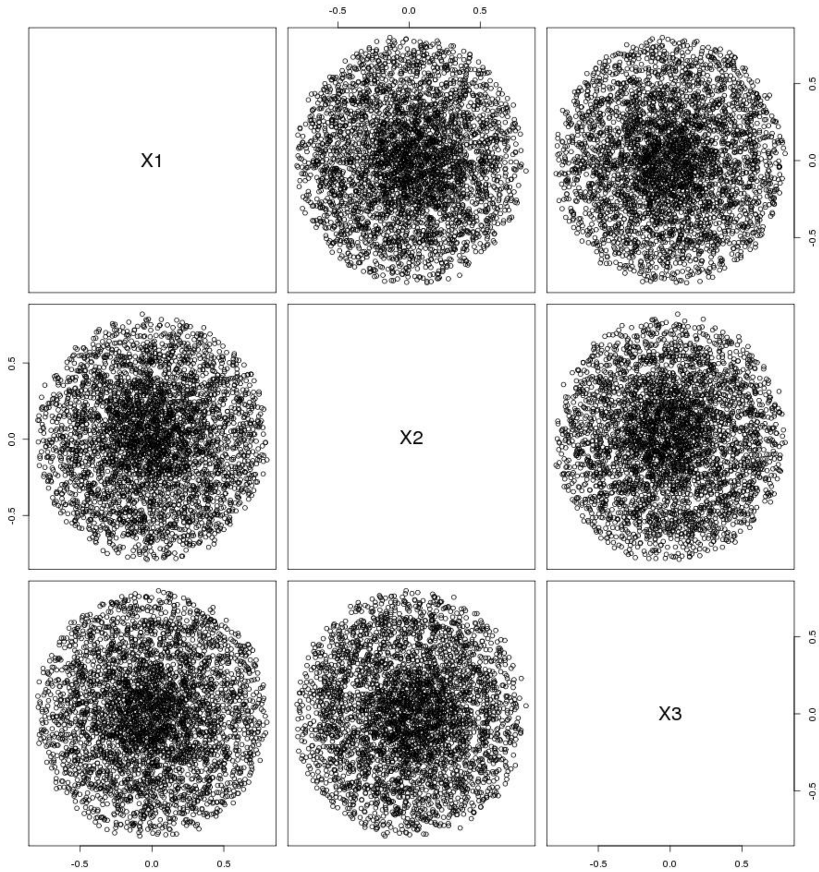
The above analyses were repeated separately for each of the transitional water types abovementioned, i.e. types D, E and F.

Regarding cluster analyses, for a p-value higher than 0.95, 463 groups of 2 to 47 samples were segregated in Type D water bodies, 2 groups of 119 and 452 samples in Type E water bodies and 308 groups of 2 to 96 samples in Type F water bodies. For a lower confidence ($p \geq 0.75$), 703 groups of 2 to 53 samples were segregated in Type D water bodies, 2 groups of 119 and 452 samples in Type E water bodies and 403 groups of 2 to 97 samples were segregated in Type E water bodies.

In general, these analyses led to an unmanageable number of groups. However, for Type E water bodies the group containing 452 samples included data from FR, SP and RoI & UK (the group containing 119 samples included data only from RoI & UK) could be further explored for IC. Nevertheless, German water bodies would be missing in such group.

Likewise, when samples were plotted in an nMDS for each of the transitional water types (i.e. D, E and F) after the Bray-Curtis distances calculated, they occupied all available 3-dimensional space, not showing any clustering (Figure 6, Figure 7, Figure 8). Moreover, the minimum stress value for the ordinations was high (0.330, 0.329 and 0.329, respectively), indicating that the points were close to being arbitrarily placed in the 3-dimensional ordination (Clarke and Warwick, 2001).

Figure 6. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type D transitional waters. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.330.



.

Figure 7. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type E transitional waters. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.329.

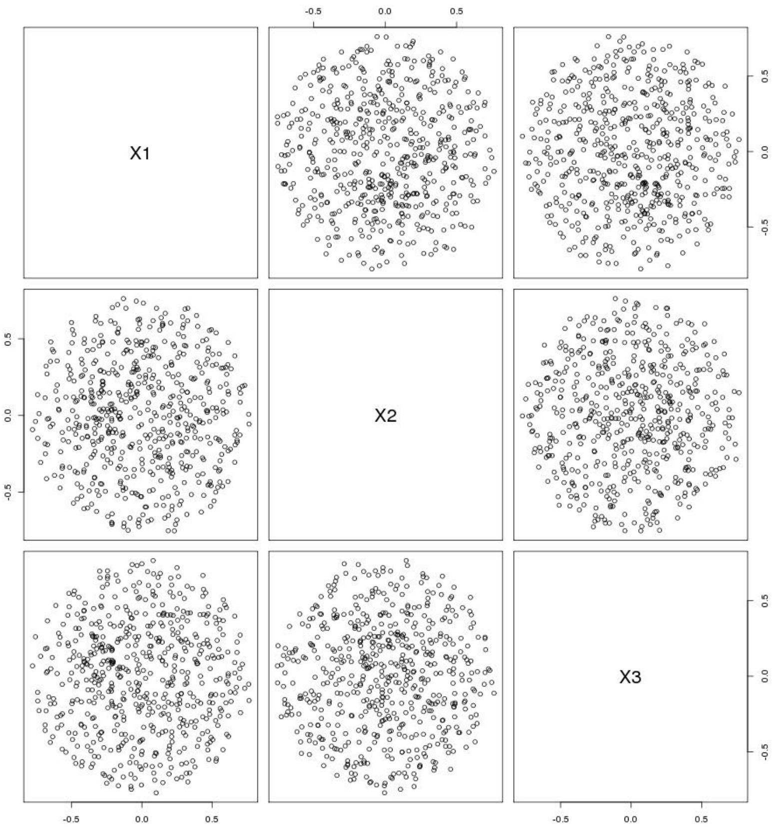
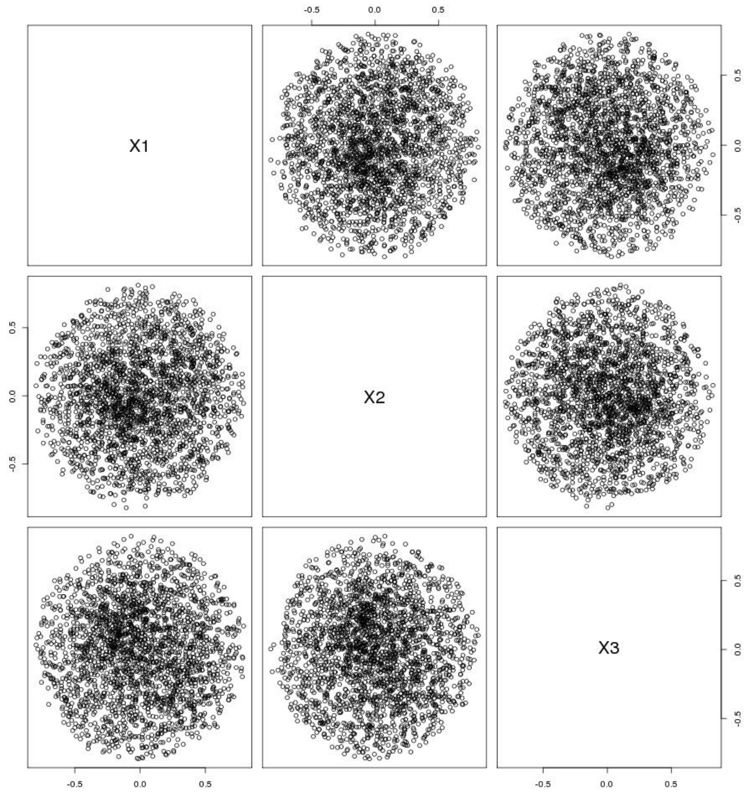


Figure 8. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type F transitional waters. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.329.



Likewise, when the samples were identified according to the saline stretch (Figure 9, Figure 10, Figure 11) or tidal level (intertidal/subtidal) (Figure 12, Figure 13, Figure 14), groups were not found. A possible exception could be found for oligohaline stretch in Type F transitional waters which seem to be more or less grouped (red samples in Figure 11), but those samples come only from two Portuguese estuaries (Mira and Mondego). So, that group is not useful for IC.

Figure 9. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type D transitional waters. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.330. Black: samples from fresh water stretches; Red: samples from oligohaline stretches; Green: samples from mesohaline stretches; Blue: samples from polyhaline stretches; Cyan: samples from euhaline water stretches.

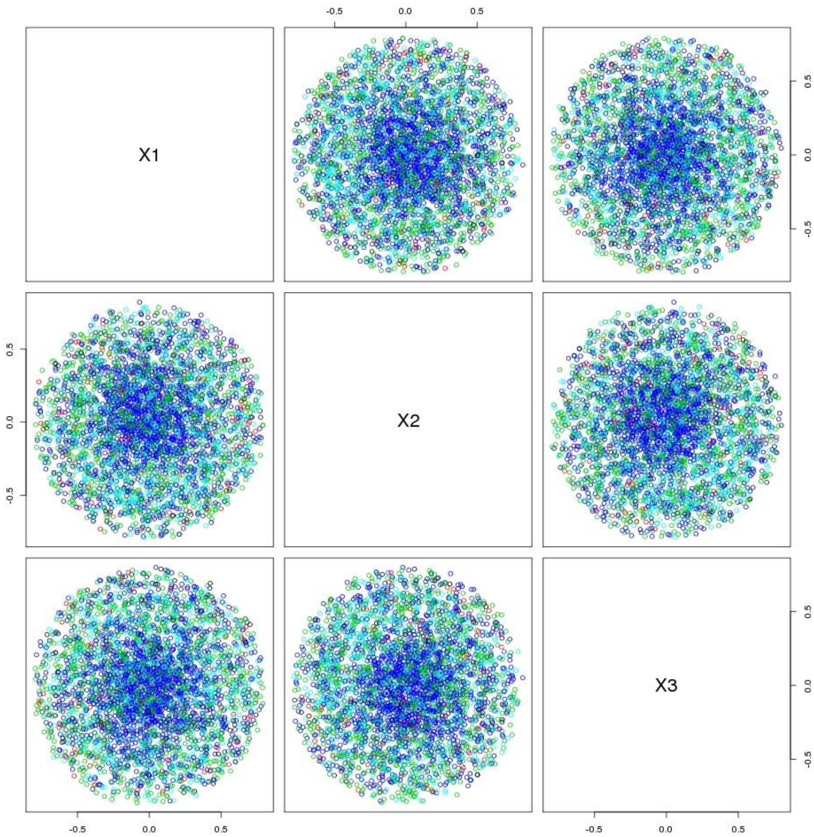


Figure 10. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type E transitional waters. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.329. Red: samples from oligohaline stretches; Green: samples from mesohaline stretches; Blue: samples from polyhaline stretches; Cyan: samples from euhaline water stretches.

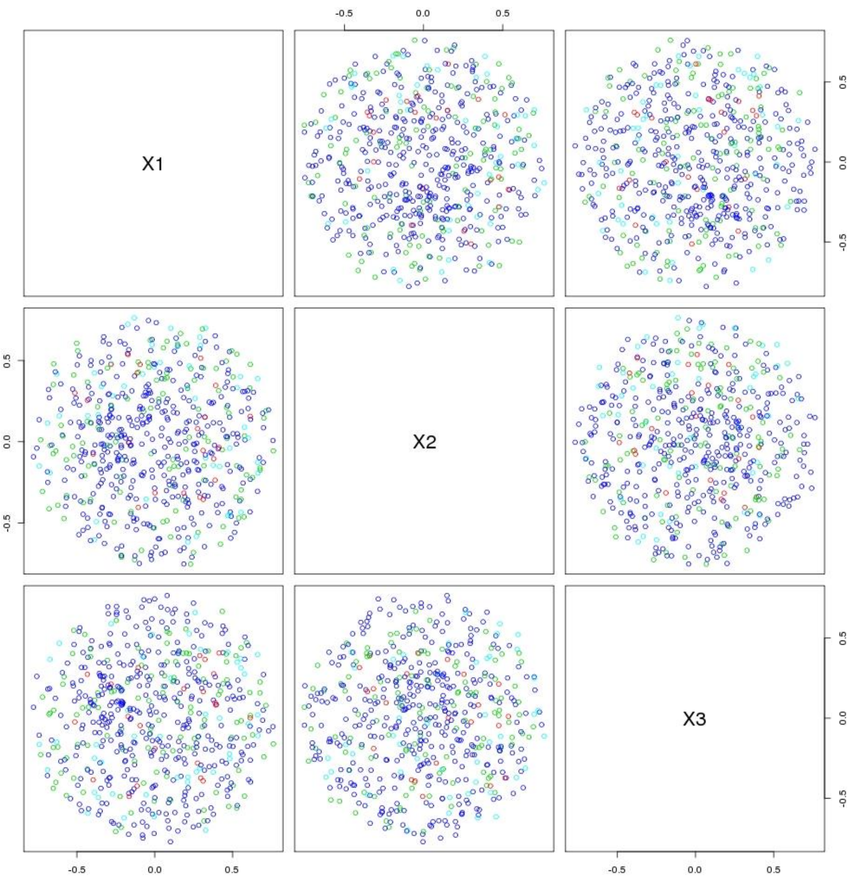


Figure 11. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type F transitional waters. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.329. Black: samples from fresh water stretches; Red: samples from oligohaline stretches; Green: samples from mesohaline stretches; Blue: samples from polyhaline stretches; Cyan: samples from euhaline water stretches.

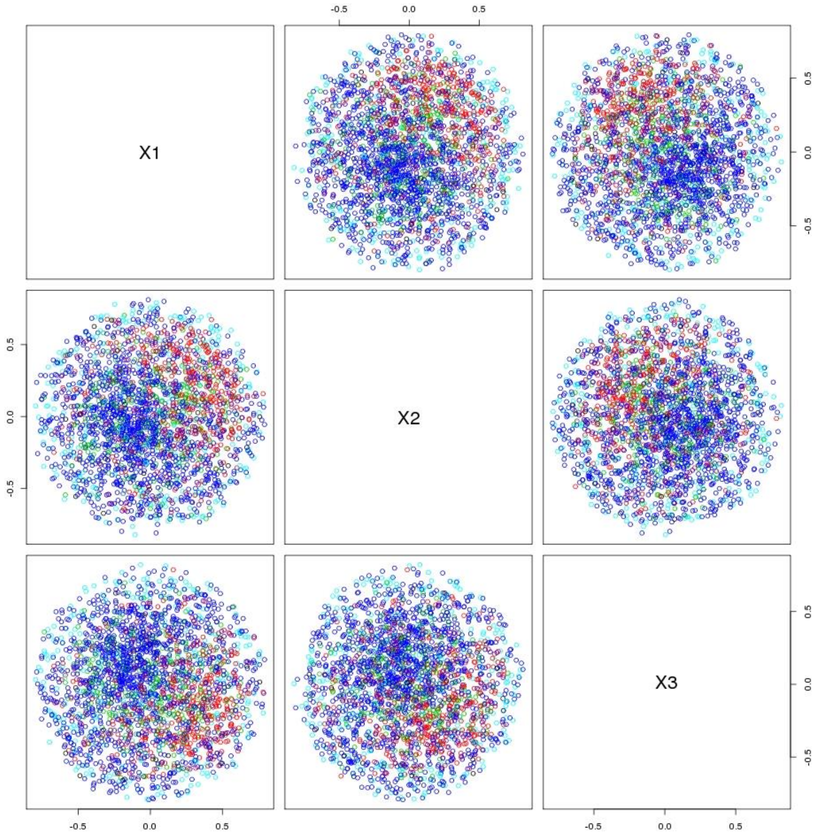


Figure 12. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type D transitional waters. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.330. Black: samples from intertidal sampling stations; Red: samples from subtidal sampling stations.

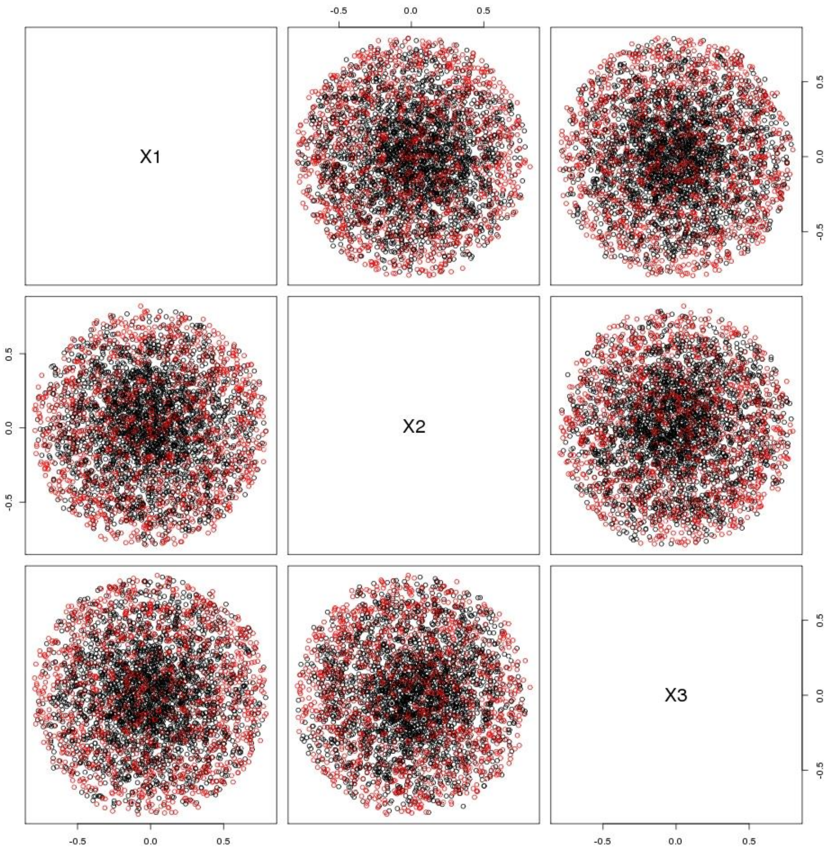


Figure 13. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type E transitional waters. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.329. Black: samples from intertidal sampling stations; Red: samples from subtidal sampling stations.

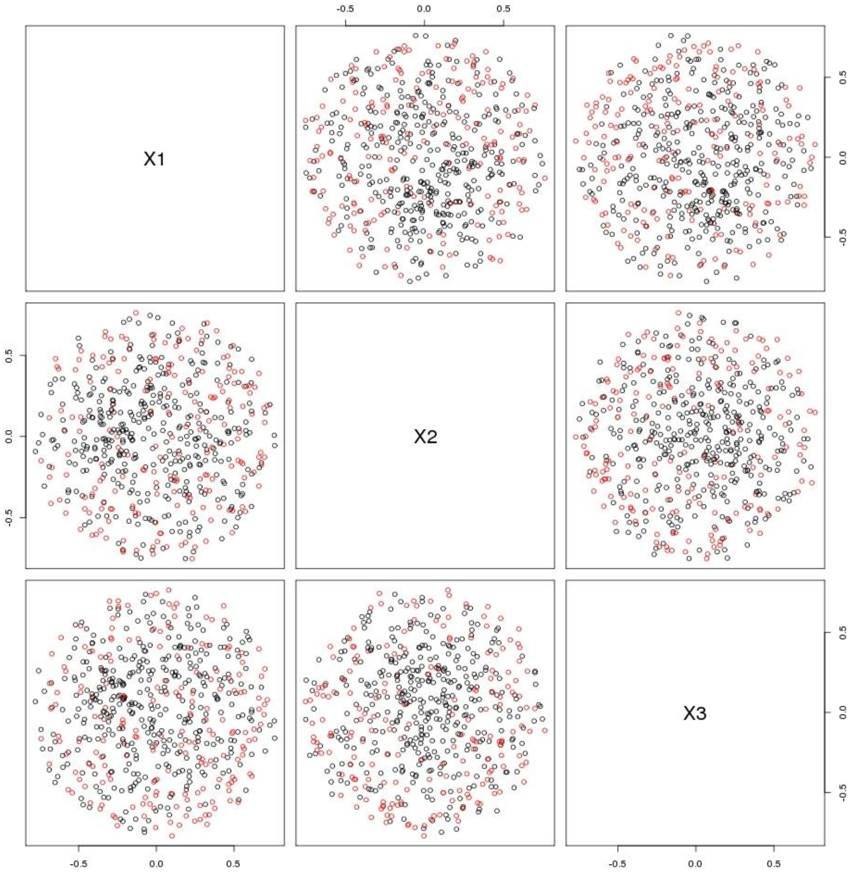


Figure 14. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type F transitional waters. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.329. Black: samples from intertidal sampling stations; Red: samples from subtidal sampling stations.

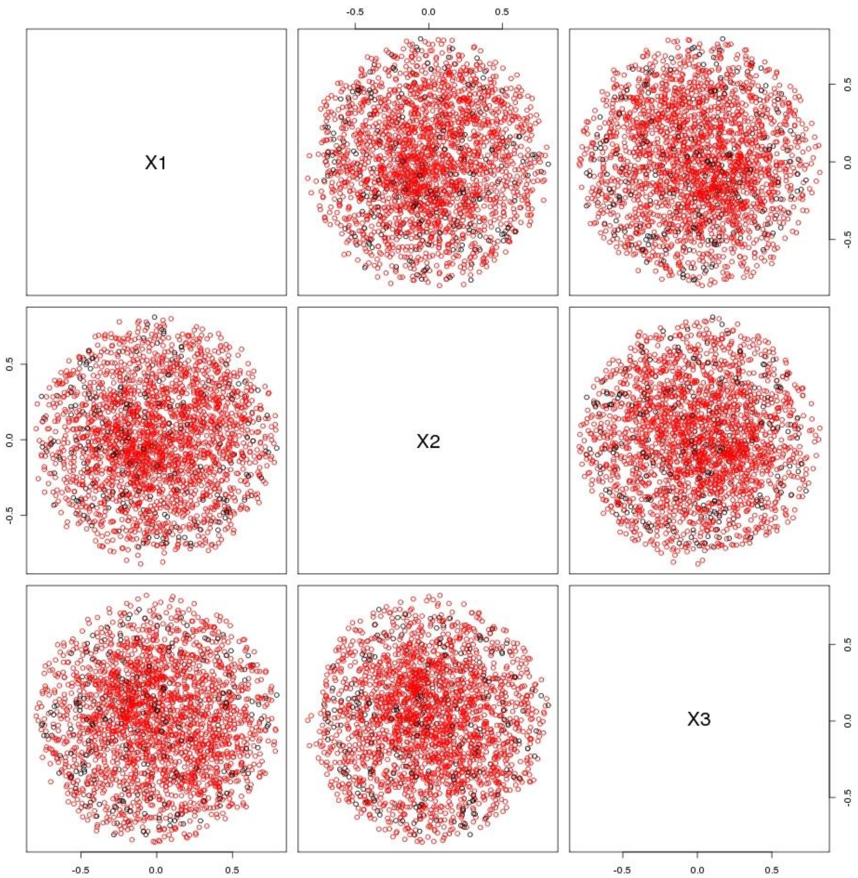


Table 10. IC types for which assessment methods are appropriate. RC: reference conditions.

Method	Appropriate for IC types / subtypes	Remarks
BAT	RC for types D and F	RC varied by ecotope within the type
BEQI2	Type D	RC estimated using 99-percentile index values from 15 year dataset
IQI	Types D, E and F	RC modelled from datasets and estimated after physicochemical conditions and sampling method
M-AMBI	RC for types D, E and F	RC varied by ecotope within the type
QSB	RC for type E	
TaSBEM	RC for types D, E and F	Best performance expected for types C and D

4.2 Pressures addressed

4.2.1 Available pressure-response information

The assessment methods respond to pressures. Such responses have been demonstrated elsewhere, being some examples summarized in Table 11 and in Figure 15 to Figure 32.

Table 11. Pressures addressed by the national methods included in the IC exercise and overview of the relationship between national methods and the pressures.

Member State	Method / Metrics tested	Pressure	Pressure indicators	Amount of data	Strength of relationship
DE	M-AMBI	Dredging and dumping	Dumping	3 sites	ANOVA; p<0.05
		General degradation Habitat destruction			
NL	BEQI2	Eutrophication	Dissolved oxygen	15 samples	r=0.59; p=0.017
			Nutrients (DIN)	15 samples	r=0.74; p=0.001
			Organic matter		
		Hydromorphological changes	Increased flow velocity due to deepening and straightening of the estuary (%area v_{max}>0.8 m·s^{-1})	4 samples (inter-/extrapolated to 15)	r=0.63; p=0.009
		Pollution	Organic compounds Micropollutants		
		Dredging and dumping	Sedimentation		
PT	AMBI	Eutrophication Dredging and dumping Hydromorphological changes Pollution Recovery	Nutrients, organic matter, dredging, sediment disposal, engineering works, organic compounds, metals and removal of pressures, assessed by a multipressure index (Borja *et al.*, 2011)	42 samples	r=0.49; p<0.001
	Margalef				r=-0.54; p<0.001
	Shannon				r=-0.75; p<0.001
	BAT				r=-0.73; p<0.001
			Nutrients, organic matter, dredging, sediment disposal, engineering works, organic compounds, metals and removal of pressures; temporal changes (Teixeira *et al.*, 2009; Neto *et al.*, 2010)	109 samples	Not applicable; see response to changes in pressures in Figure 20
SP-An	TaSBeM	Eutrophication Point source and diffuse source pollution	Nutrients, organic matter (assessed by t-LUSI)	34 samples	r=-0.85; p=0.003
				120 samples	r=-0.52; p=0.000
SP-BC	AMBI	Sediment chemical quality (Muxika *et al.*, 2012)	Organic matter Metals Organic compounds	372 samples	ρ=0.32; p<0.01
		Multiple pressures (Borja *et al.*, 2011)	Multipressure index	42 samples	r=0.49; p<0.001
	Richness				r=-0.37; p=0.01
	Shannon				r=-0.75; p<0.001
	M-AMBI				r=-0.72; P<0.001

		Eutrophication (Borja and Tunberg, 2011)	Dissolved oxygen Nutrients Organic matter	240 samples	r=0.61; p=0.017
		Aquaculture (Callier *et al.*, 2009)	Finfish and shellfish production Organic matter	7 sites	r^2=0.68; p=0.022
		Dredging and dumping (Borja et al., 2009)	Dredging	5 samples	t-test (before/after comparison) p<0.005
				10 samples	t-test (before/after comparison) p<0.005
				6 samples	t-test (before/after comparison) p<0.05
			Sediment disposal	9 samples	t-test (before/after comparison) p<0.005
				3 samples	t-test (before/after comparison) p<0.005
		Hydromorphological changes (Borja *et al.*, 2009)	Land reclamation	5 samples	t-test (before/after comparison) p<0.005
		Recovery (Borja *et al.*, 2009)	Discharge removal	12 samples	t-test (before/after comparison) p<0.005
				7 samples	t-test (before/after comparison) p<0.005
			Oxygen saturation	49 samples	r=0.85; p<0.001
SP-C	QSB	Eutrophication	Oxygen Organic matter	50 samples (Puente *et al.*, 2010)	r=0.64; p<0.01
		Pollution	Metals Organic compounds		
		Engineering works	Hydromorphological changes		
RoI & UK	IQI	Eutrophication	Organic carbon	176 samples	r^2=0.71; p<0.001
		Pollution	Metals		
			Particulates	213 samples	r^2=0.34; p<0.001
		Dredging and dumping	Aggregate extraction, Dumping		
		General disturbance			

Figure 15. Comparison of M-AMBI EQR values direct in a dumpsite for dredged material (Dumping), in an adjacent area (drift) and in a reference area in the river Weser.

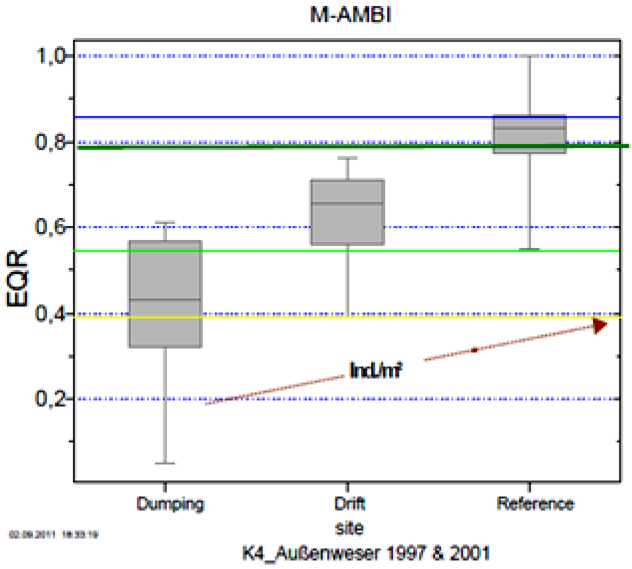


Figure 16. Correlation between dissolved oxygen (measured at Schaar van Ouden Doel) and benthos EQR (BEQI2) for the ecotope Mesohaline-Intertidal. This analysis was initially made using 99/1 percentile reference values.

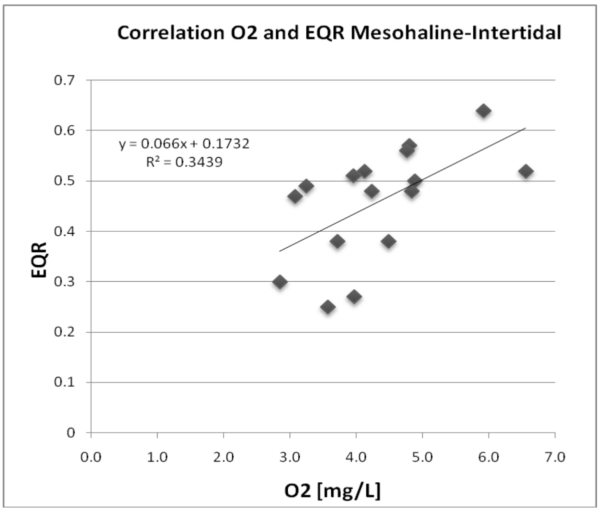


Figure 17. Correlation between DIN (measured at Schaar van Ouden Doel) and benthos EQR (BEQI2) for the ecotope mesohaline-intertidal.

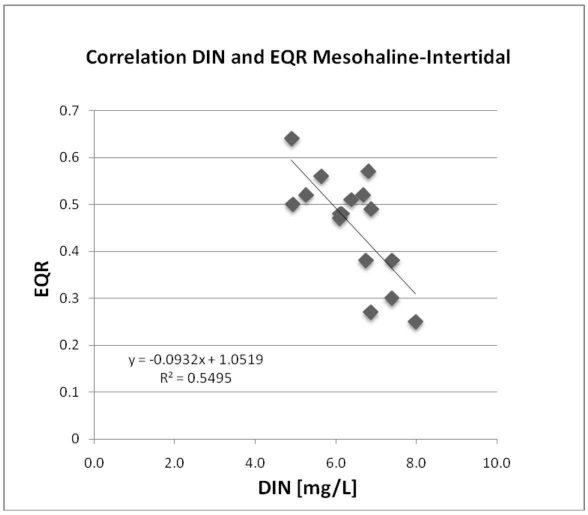


Figure 18. Correlation between the percentage of high-dynamic litoral ecotope in the Westerschelde (mostly polyhaline ecotope; derived from Van den Bergh et al.*, 2003, figure 2.4) and the EQR (BEQI2) of the polyhaline-subtidal ecotope.*

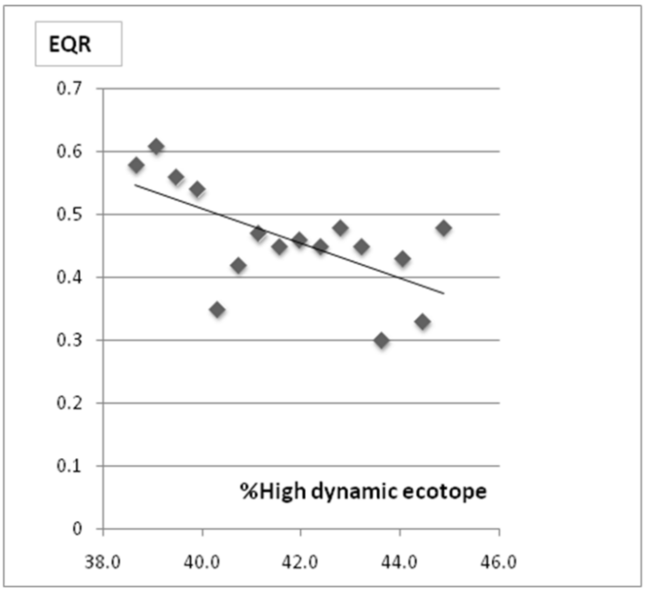


Figure 19. Linear regression between multipressure index and BAT (Borja et al.*, 2011).*

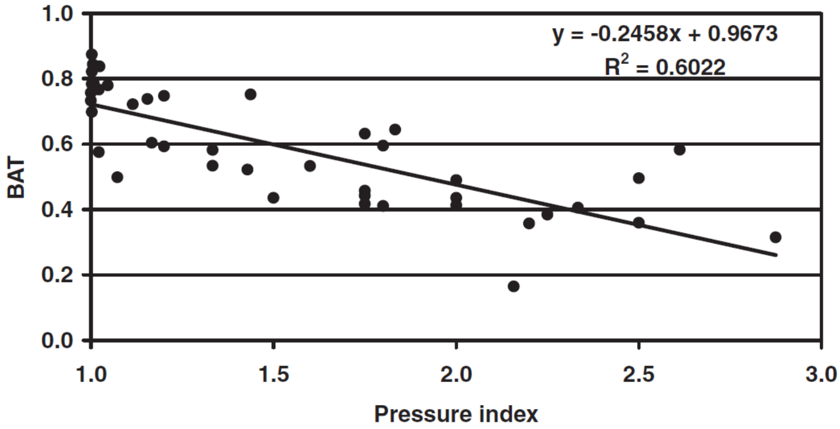


Figure 20. EQR and EQS variation given by BAT method in the period from 1990 to 2006 at the Mondego estuary, and table including the events to which the changes in BAT should be related.

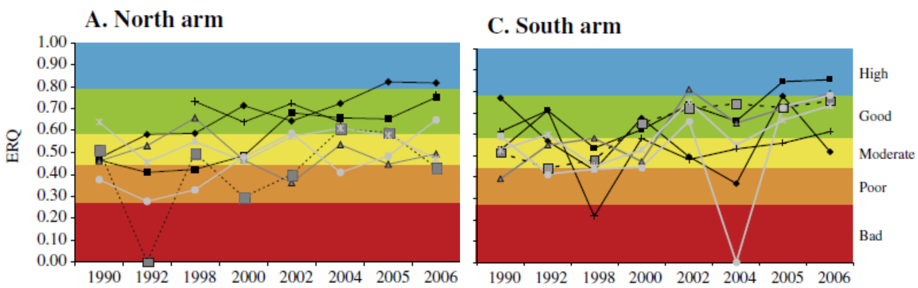


Table 1 List of events regarding Mondego estuary morphological evolution from 1791 to 2006

Events	References	Duration	Location	Effects on the system
First regularization of Lower Mondego valley between Coimbra and the Foja river	AGSHE 1934	1791–1807	Lower Mondego	Margins regularization; Drainage of Lower Mondego valley fields
Redefinition of external physical limits of the estuary (external margins, docks, wave breaks)	Proença Cunha et al. 1997; Marques et al. 2005; 2007.	1968	Stations 1, 2, 3	River bed dredged and deepened near the mouth; Embankment and consolidation of margins
Construction of internal harbor physical limits (docks, pears, mercantile harbor)	Proença Cunha et al. 1997; Marques et al. 2005; 2007	1980s	Stations 1, 2, 3 and 10	River bed dredged and deepened near the mouth; Embankment and consolidation of margins; Construction of harbor's land infrastructures
Regularization of the river bed, through the lower Mondego area until upstream of the separation point	Proença Cunha et al. 1997; Marques et al. 2005; 2007	Second half of 1980s	Upstream the study area	Reduction of intertidal areas (wetlands) at north arm; River bed deepened and turned slender; Embankment and consolidation of margins
Regularization of river bed between separation point and the harbor (promoted in 1992 the complete interruption between north and south arms)	Proença Cunha et al. 1997; Marques et al. 2005; 2007; Teixeira et al. 2009	1990s	Stations 10 to 14	Reduction intertidal mudflat areas, embankment and consolidation of margins at north arm; Macroalgal blooms and *Zostera noltii* coverage reduction due to environmental degradation at south arm
Interruption between north and south arm kept on	Martins et al. 2001; Marques et al. 1997; 2003; 2005; 2006; 2007; Teixeira et al. 2009	1993 to 1997	Southern arm	Residence time and nutrient loading increased in the south arm; Macroalgal blooms came frequent and *Zostera noltii* coverage reduction in the south arm
Construction of solid gross materials port at the mercantile harbor	Proença Cunha et al. 1997; Marques et al. 2005;2007	1997	Between stations 2 and 10	Disturbance of bottoms at the final part of north arm
Implementation of mitigation measures: 1) Reestablishment of communication (1 m^2) between north and south arms; 2) Reduction of nutrient inputs from Pranto river sluice discharge	Neto 2004; Lillebø et al. 2005; Neto et al. 2008; Marques et al. 2005; 2007; Teixeira et al. 2009	1997/98	1) Station 14; 2) South arm	Residence time and nutrient loadings decreased at south arm; Macroalgal blooms ceased and *Zostera noltii* coverage start recovering at south arm
Centenary floods	Teixeira et al. 2009	Winter 2000/01		Current increased at north arm
Waste water treatment plant (Fontela)	Teixeira et al. 2009	2003	Station 12	Disturbance of bottoms
Construction of a small boats fishing harbor	Teixeira et al. 2009	June to October 2004	Station 3	Reduction of intertidal areas at south arm; Embankment and consolidation of margins
Severe drought	Marques et al. 2006; 2007; Teixeira et al. 2009	2005		Reduction of water flow and freshwater influence on lower areas of the estuary
Reestablishment of communication between both arms	Marques et al. 2005; 2006; 2007; Teixeira et al. 2009	Spring 2006	Station 14	Reduction of residence time at south arm

Figure 21. Scatterplot for TaSBeM and pressure indicators (sediment soluble phosphorous (P) and transitional-LUSI (tLUSI_d)) in transitional Atlantic waters of Andalusia, using monitoring data. The relationship between TaSBeM and P should not be taken into account since P was used to derive the linear combination of component metrics in TaSBeM.

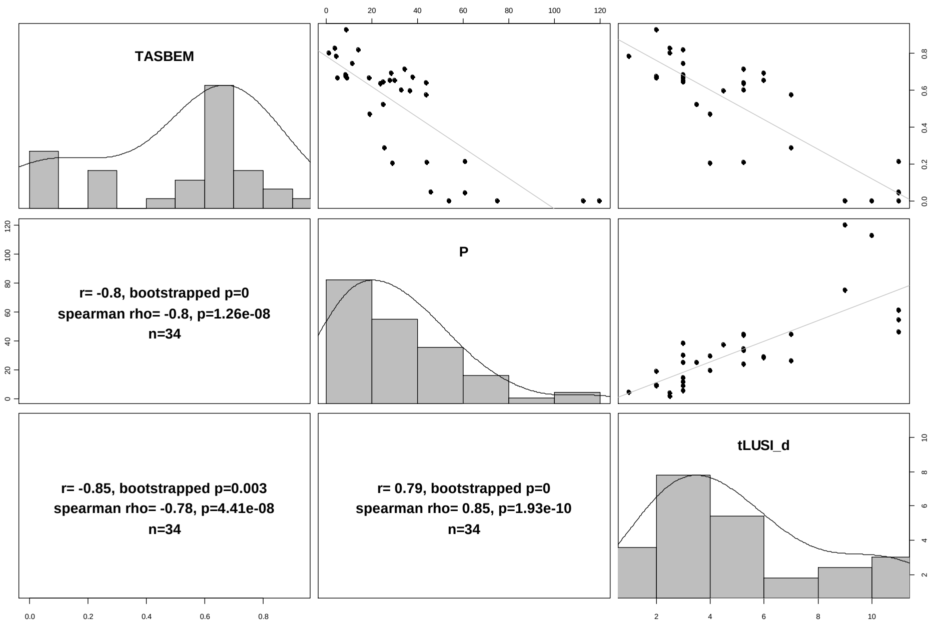


Figure 22. Scatterplot matrix for TaSBeM and transitional-LUSI (tLUSI) in transitional Atlantic waters of Andalusia using historical data.

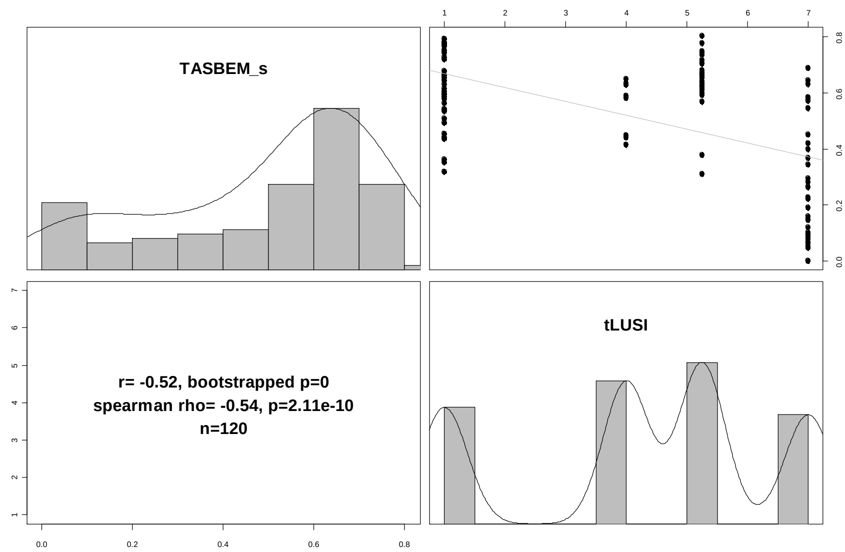


Figure 23. Linear regression between pressure index (Borja et al.*, 2011) and M-AMBI.*

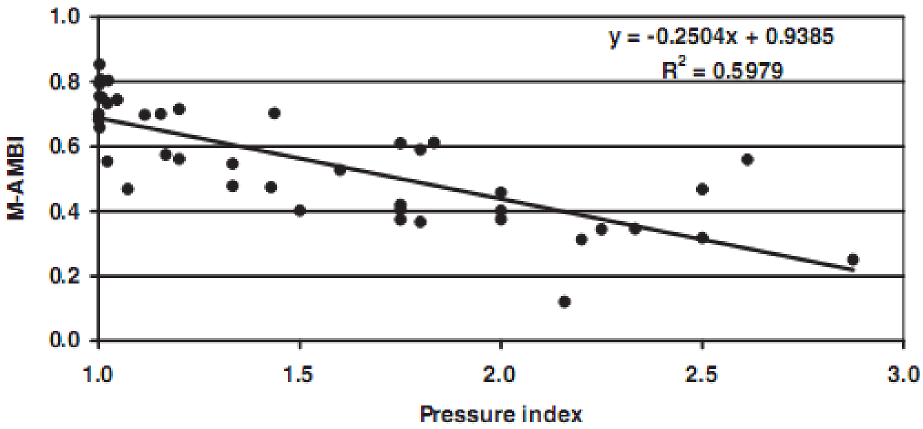


Figure 24. Linear relationship between density of mussels and M-AMBI (Callier et al.*, 2011).*

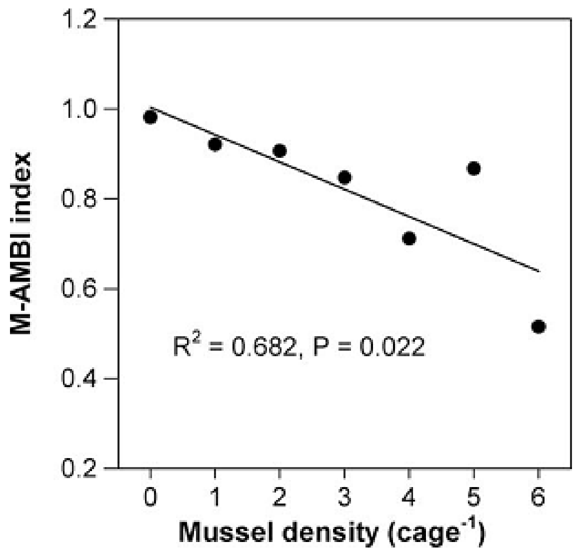


Figure 25. Table summarizing the response of M-AMBI to changes in pressures (discharge removal, dredging, sediment disposal, land reclamation and marina construction) (Borja et al.*, 2009).*

water body	stations	pressure/ action	before pressure or action: years	before pressure or action: mean M-AMBI ± SD	after pressure or action: years	after pressure or action: mean M-AMBI ± SD	Student's t	P significance
Mompás–Pasaia	41	discharge removal	1996–2001	0.34 ± 0.13	2002–2007	0.85 ± 0.13	−6.56	<0.005
Getaria–Higer	45	dredging disposal	1995, 1997–2000	0.80 ± 0.04	1996, 2001–2003	0.53 ± 0.11	5.53	<0.005
	47	dredging disposal	2002	0.80 ± 0.04	2003 and 2004	0.53 ± 0.11	5.53	<0.005
Butroe	12	discharge removal	1995–1997	0.66 ± 0.02	1998–2001	0.79 ± 0.03	−7.84	<0.005
Butroe	12	dredging	1999–2001	0.81 ± 0.02	2002	0.50 ± 0.06	8.50	<0.005
	11	dredging		0.81 ± 0.02	2002 and 2003	0.50 ± 0.06	8.50	<0.005
Oka	17	dredging	1997, 1998, 2001, 2002, 2005	0.55 ± 0.07	1995, 1999, 2000, 2003, 2004	0.39 ± 0.04	4.28	<0.005
Outer Nervión	7	dredging	1998–2001	0.84 ± 0.16	2002 and 2003	0.38 ± 0.16	3.44	<0.05
Orio	35	land reclamation	1999–2001	0.48 ± 0.02	2002 and 2003	0.32 ± 0.06	3.93	<0.005
	35	marina construction	2004 and 2005	0.59 ± 0.07	2006 and 2007	0.48 ± 0.07	1.63	n.s.

The years used within the analysis are shown for each station and water body.
SD, standard deviation; n.s., correlation not significant (P > 0.05).

Figure 26. Exponential regression between M-AMBI and oxygen saturation in samples from estuaries submitted to recovery schemes.

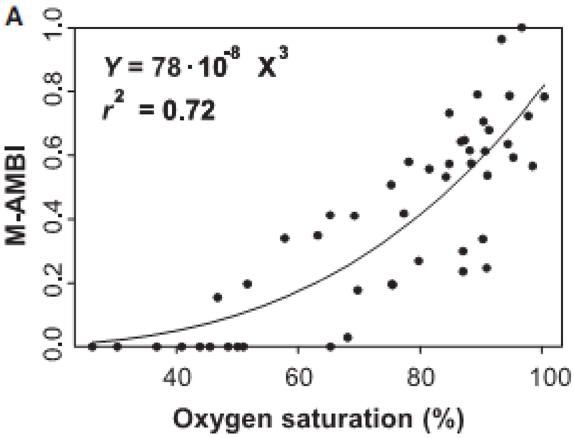


Figure 27. Spearman correlation between the QSB index and the index of global pressure.

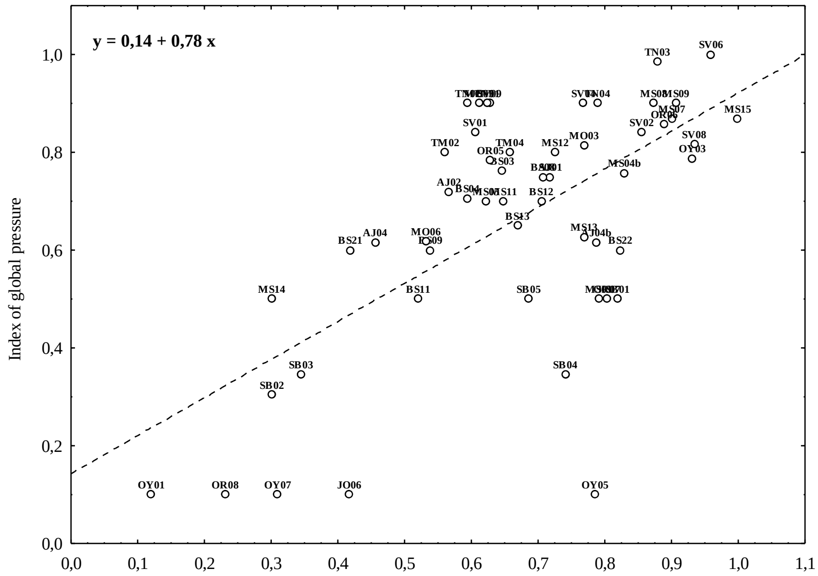


Figure 28. Normalised IQI values versus summed pressure scores for Basque Coast with Pearson correlation (data from Borja et al.*, 2011).*

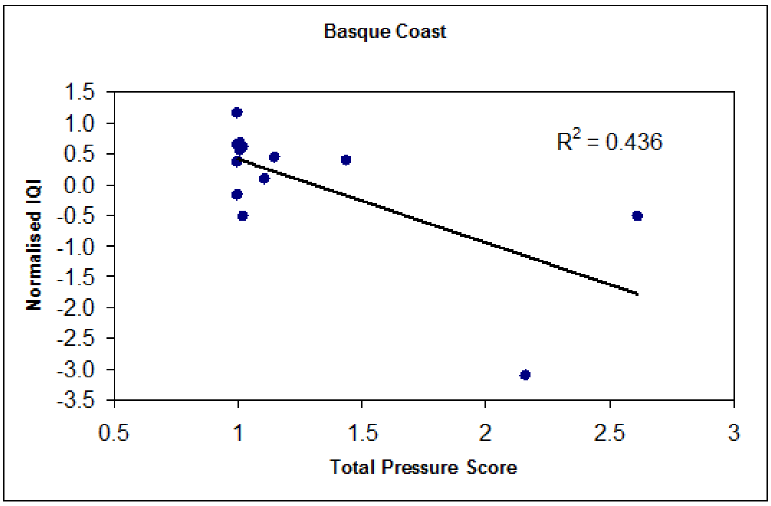


Figure 29. Normalised IQI values versus summed pressure scores for Lesina Lagoon with Pearson correlation (data from Borja et al.*, 2011).*

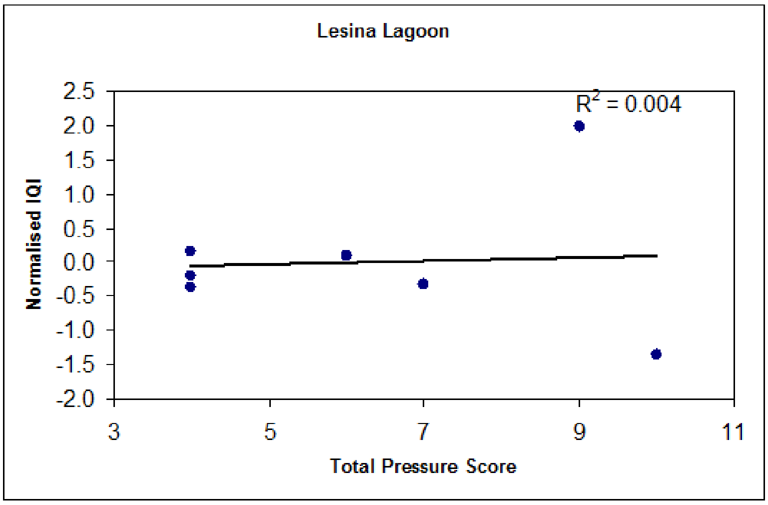


Figure 30. Normalised IQI values versus summed pressure scores for Mondego Estuary with Pearson correlation (data from Borja et al.*, 2011).*

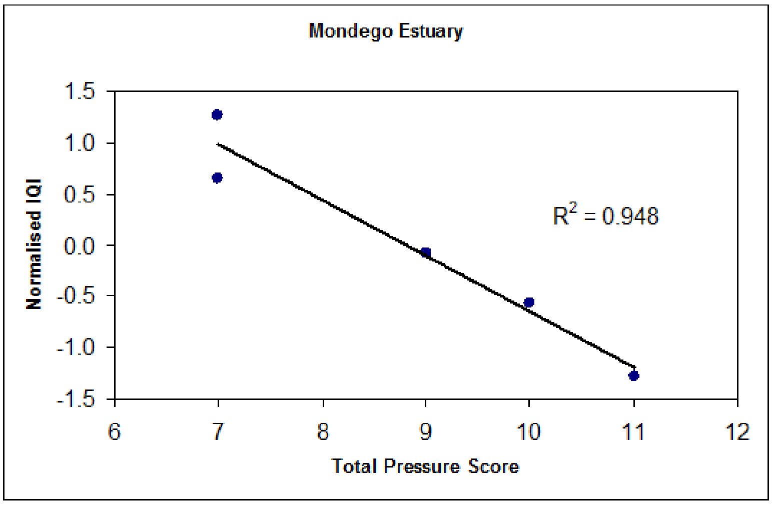


Figure 31. Normalised IQI values versus summed pressure scores for Oslofjord with Pearson correlation (data from Borja et al.*, 2011).*

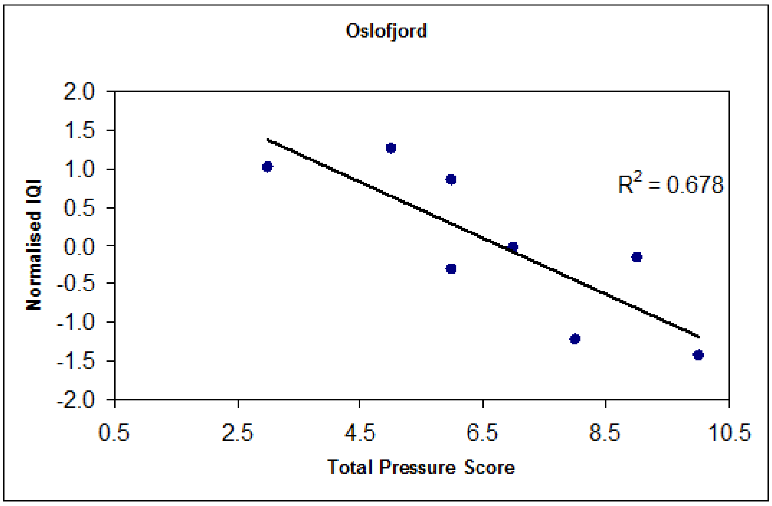
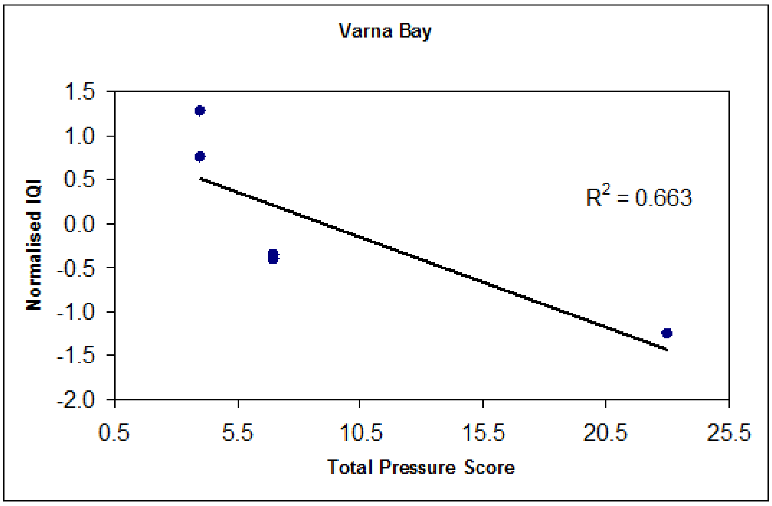


Figure 32. Normalised IQI values versus summed pressure scores for Varna Bay with Pearson correlation (data from Borja et al.*, 2011).*



4.2.2 Pressure-response relations within the common dataset

STEP1: Including different pressures.

Being the quality of pressure data available in the IC database adequate, most of the significant relationships described above should be found also when testing the EQR values in the common dataset against the pressure values. However, the pressures measured for the exercises above and the pressures included in the database are not exactly the same. Hence, some assumptions had to be made:

- The total pressure included in the database has been considered as a proxy for several multipressure indices used in the examples above.
- The hydromorphological changes in the database have been considered as a proxy of hydromorphological impacts mentioned in the examples above.

Moreover, some of the pressures could not be taken into account as any of the pressures included in the database seemed to address them. That is the case, e.g., for eutrophication (and measures related with it, i.e. dissolved oxygen and inorganic nitrogen, and t-LUSI) or organic enrichment.

This exercise was undertaken in a first phase including data from Type D water bodies. Taking into account the available data: BEQI2 was tested against hydromorphological pressures; BAT (using national reference conditions) was tested against hydromorphological pressures and total pressure; M-AMBI was tested against total pressure, hydromorphological pressure, sediment chemical quality, capital dredging and maintenance dredge disposal volume; and IQI was tested against sediment chemical quality and total pressure.

Regressions were significant ($p<0.01$) for all analyses except for BEQI2 and M-AMBI against hydromorphological pressure ($p=0.731$ and $p=0.122$, respectively) (Table 12). However, all significant regression lines had positive slopes, i.e. EQR values were higher in samples with higher pressure level, which is just the contrary of what it should be expected. The only exceptions were those of IQI and M-AMBI against sediment chemical quality, with a negative slope. Nevertheless, residuals were not normally distributed, which is one of the requirements of lineal regression analysis.

Table 12. Results from the regressions for methods against pressures. Tot Pre = Total Pressure index; Hyd Cha = Hydromorphological Changes; Sed Che = Sediment Chemical Quality; Cap Dre = Capital Dredging; Dis Vol = Maintenance Disposal Volume; IC RC = reference conditions defined by the intercalibration group; nat RC = reference conditions defined by Portuguese experts; r^2 = coefficient of determination.

		Tot Pre	**Hyd Cha**	**Sed Che**	**Cap Dre**	**Dre Dis**
BEQI2	**slope**	-	0.000	-	-	-
	r^2	-	0.004	-	-	-
	p	-	0.731	-	-	-
BAT	**slope**	0.001	0.003	-	-	-
	r^2	0.007	0.005	-	-	-
	p	0.003	0.020	-	-	-
M-AMBI	**slope**	0.001	-0.001	-0.015	0.004	0.008
	r^2	0.006	0.001	0.039	0.003	0.013
	p	0.000	0.122	0.000	0.001	0.000
IQI	**slope**	-	-	-0.015	-	-
	r^2	-	-	0.067	-	-
	p	-	-	0.000	-	-

These results indicated that, probably, pressures are not quantified in a useful way for this exercise. However, (1) as the only pressure value, from the above tested, to which methods responded was sediment chemical quality, (2) as the p-values indicated very high significance ($p<0.001$) and (3) as those results were supported by literature, it was assumed that both IQI and M-AMBI responded to sediment chemical quality. As the best option to find a common relation was the use of sediment chemical quality data, the remainder methods (BAT, BEQI2 and TAsBeM) were also tested.

The results showed that BAT and TAsBeM did not respond to sediment chemical quality (regressions were not significant), whereas regressions were significant for BEQI2 against sediment chemical quality (Table 13). Moreover, this regression line had negative slope, which means that, as expected, EQR values were lower for higher pressure levels. However, as in the abovementioned cases, residuals were not normally distributed, which is one of the requirements of lineal regression analysis.

Table 13. Results from the regressions for methods against sediment chemical quality (Sed Che). IC RC = reference conditions defined by the intercalibration group; nat RC = reference conditions defined by Portuguese experts; r^2 = coefficient of determination.

		Sed Che
BEQI2	**slope**	-0.011
	r^2	0.030
	p	0.000
BAT	**slope**	-0.003
	r^2	0.001
	p	0.285
TAsBeM	**slope**	0.000
	r^2	0.000
	p	0.815

These results have been plotted in Figure 33. From the plots, it can be seen that there is a lower dispersion of EQR values (for each pressure value) for IQI, which could explain why IQI performs better that the remainder of methods. In turn, such a lower dispersion would be due, at least in part, to the fact that there are fewer samples with IQI data.

In Figure 34 all the regression lines have been plotted together in order to compare them. From that figure, it can be seen that IQI and M-AMBI respond to sediment chemical quality in the same way (parallel lines). However, IQI values are 0.15 units higher than M-AMBI values for the same pressure value. Conversely, BAT and BEQI2 respond to sediment chemical quality also in a similar way, but being BAT values 0.10 units higher than BEQI2 values for the same pressure level. Finally, TAsBeM does not respond to changes in the sediment chemical quality.

Regression lines for methods against sediment chemical quality were all significant, except for TAsBeM. However, slopes were very low. Moreover, as abovementioned, residuals were not normally distributed, which is one of the requirements of lineal regression analysis.

As at least some of the methods were tested against metal concentrations (IQI and M-AMBI) in literature, and metal concentrations are taken into account to assess the sediment chemical quality level in the database, it is concluded that the quality of the pressure measures included in the database is not sufficient.

Figure 33. Regression lines from each of the assessed methods (Y-axis) against sediment chemical quality (X-axis). Colours indicate the country each of the samples come from.

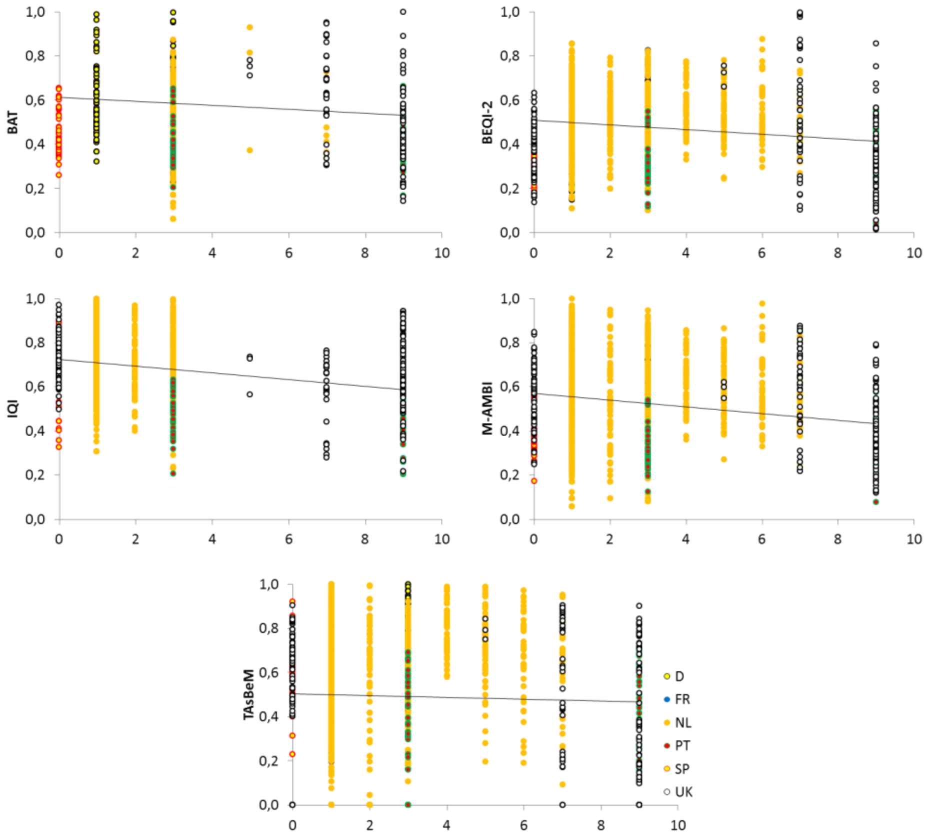
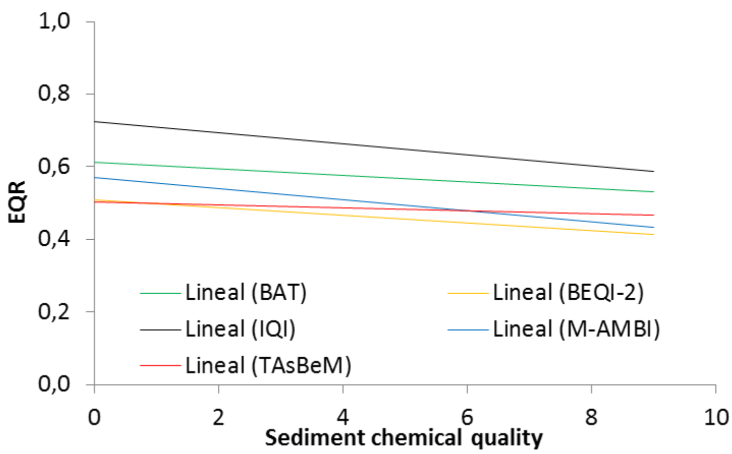


Figure 34. Comparison of regression lines for each of the methods against sediment chemical quality using all the available data in the database.



One of the problems, already highlighted in the 2nd phase of IC, could be that, most member states assessed the pressure level at water body level or salinity stretch level. Moreover, there are pressure data from single years which are repeated for sampling sites monitored in different years (same pressure value and different EQR values). This would let to high variability in EQR values for different levels of pressure and to poor correlations between the metrics and the pressure values.

With available data, that problem seemed not easy to solve. A solution would be to ask member states to select, from the database, a subset of samples with clear pressure gradients or trends and with quantitative pressure data. Those subsets would consist on: samples taken at a single site, where there is a known trend in pressure, in different years; or samples taken at sites along a known pressure gradient.

STEP 2: Sediment chemical quality data: sample level

Such information was requested and the database was reduced to 203 samples from DE, FR, NL, SP and UK, including the information summarized in Table 14.

Table 14. Number of pollutant data (and environmental data) provided by each of the Member States and total number.

	DE	FR	NL	SP-BC	SP-C	UK	TOTAL
Salinity	18	65	19	11	0	39	152
O_2	0	0	0	11	50	30	91
%Mud (Φ<20 µm)	18	0	0	0	0	0	18
%Mud	18	65	0	11	48	39	181
%Organic Matter	18	65	0	11	50	39	183
Total Organic C	0	65	0	0	0	0	65
REDOX potential	18	0	0	10	0	0	67
Al	0	65	0	0	0	0	65
As	18	28	20	8	50	39	163
Cd	18	65	20	10	50	39	202
Cr	18	53	20	10	50	39	190
Cu	18	53	20	10	50	39	190
Hg	18	65	20	10	50	39	202
Ni	18	65	20	10	50	39	202
Mn	0	25	20	10	0	39	94
Fe	0	25	20	10	50	39	144
Pb	18	65	20	10	50	39	202
Zn	18	53	20	10	50	39	190
PAH	18	53	20	10	10	39	150
PCB	18	53	20	10	0	39	140

Correlation analyses between environmental and contaminants data were carried out, and also linear regressions between methods, and environmental and contaminants data in order to check the response of the methods to pressures. Such information has been summarized in Table 15 and

Table 15. Correlations between environmental and contaminant data, from the reduced dataset. r: Pearson; p: significance (p-value); N: number of pairs of data included in the analysis. Significant correlations (p<0.05) with 0.25≤|r|<0.50 are shaded in orange; significant correlations (p<0.05) with 0.50≤|r|<0.75 are shaded in green; Significant correlations (p<0.05) with 0.75≤|r| are shaded in blue; Non-significant correlations (p>0.05) are striped.

		Salinity	O_2	%Mud (Φ<20 µm)	%Mud	%Organic Matter	Total Organic C	REDOX potential	Al	As	Cd	Cr	Cu	Hg	Ni	Mn	Fe	Pb	Zn	PAH
O_2	r	-0.279																		
	p	0.077																		
	N	41																		
Mud<20	r	0.000																		
	p	1.000																		
	N	18																		
Mud	r	0.151	-0.217*	0.500*																
	p	0.083	0.041	0.035																
	N	133	89	18																
OM	r	0.304**	-0.219*	1.000**	0.500**															
	p	0.000	0.037	.	0.000															
	N	133	91	18	181															
TOC	r	-0.267*			0.426**	0.406**														
	p	0.032			0.000	0.001														
	N	65			65	65														
REDOX	r	0.152	-0.115	-0.500*	-0.584**	-0.352														
	p	0.440	0.751	0.035	0.001	0.066														
	N	28	10	18	28	28														
Al	r	-0.741**			0.586**	0.564**	0.770**													
	p	0.000			0.000	0.000	0.000													
	N	65			65	65	65													
As	r	-0.120	-0.132	0.500*	0.202*	0.167*	0.498**	-0.582**	1.000**											
	p	0.207	0.220	0.035	0.016	0.046	0.007	0.002	.											
	N	112	88	18	141	143	28	25	28											
Cd	r	-0.161*	-0.275**	0.500*	0.133	0.482**	0.571**	0.136	0.729**	0.538**										
	p	0.049	0.009	0.035	0.075	0.000	0.000	0.499	0.000	0.000										
	N	151	90	18	180	182	65	27	65	163										
Cr	r	-0.293**	-0.013	-1.000**	0.348**	0.150	0.850**	0.417*	0.978**	0.748**	0.510**									
	p	0.000	0.904	.	0.000	0.052	0.000	0.031	0.000	0.000	0.000									
	N	139	90	18	168	170	53	27	53	163	190									
Cu	r	0.070	-0.097	0.500*	0.337**	0.361**	0.919**	0.030	0.984**	0.781**	0.645**	0.786**								
	p	0.411	0.365	0.035	0.000	0.000	0.000	0.882	0.000	0.000	0.000	0.000								
	N	139	90	18	168	170	53	27	53	163	190	190								
Hg	r	0.067	-0.218*	0.500*	0.238**	0.400**	0.735**	0.224	0.704**	0.602**	0.727**	0.438**	0.724**							
	p	0.415	0.039	0.035	0.001	0.000	0.000	0.262	0.000	0.000	0.000	0.000	0.000							
	N	151	90	18	180	182	65	27	65	163	202	190	190							

Ni	**r**	-0.204*	-0.018	0.500*	0.455**	0.313**	0.801**	0.285	0.984**	0.703**	0.518**	0.795**	0.890**	0.610**						
	p	0.012	0.864	0.035	0.000	0.000	0.000	0.149	0.000	0.000	0.000	0.000	0.000	0.000						
	N	151	90	18	180	182	65	27	65	163	202	190	190	202						
Mn	**r**	-0.282**	0.258		0.293*	0.592**	0.250	-0.517	1.000**	0.691**	0.785**	0.951**	0.788**	0.561**	0.824**					
	p	0.006	0.108		0.011	0.000	0.228	0.154	.	0.000	0.000	0.000	0.000	0.000	0.000					
	N	93	40		74	74	25	9	25	67	94	94	94	94	94					
Fe	**r**	0.161	0.037		0.318**	0.363**	1.000**	0.133	0.250	0.576**	0.341**	0.439**	0.609**	0.565**	0.690**	0.448**				
	p	0.123	0.730		0.000	0.000	.	0.732	0.228	0.000	0.000	0.000	0.000	0.000	0.000	0.000				
	N	93	90		122	124	25	9	25	117	144	144	144	144	144	94				
Pb	**r**	0.045	-0.414**	-0.500*	0.440**	0.449**	0.816**	0.301	0.925**	0.746**	0.715**	0.677**	0.915**	0.723**	0.832**	0.674**	0.575**			
	p	0.583	0.000	0.035	0.000	0.000	0.000	0.128	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000			
	N	151	90	18	180	182	65	27	65	163	202	190	190	202	202	94	144			
Zn	**r**	-0.149	-0.444**	0.000	0.347**	0.546**	0.873**	0.133	1.000**	0.587**	0.847**	0.597**	0.779**	0.642**	0.719**	0.775**	0.465**	0.882**		
	p	0.080	0.000	1.000	0.000	0.000	0.000	0.509	.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		
	N	139	90	18	168	170	53	27	53	163	190	190	190	190	190	94	144	190		
PAH	**r**	0.472**	-0.194	0.866**	0.327**	0.385**	0.799**	-0.430*	0.768**	0.407**	-0.043	0.208*	0.495**	0.284**	0.487**	0.295**	0.728**	0.464**	0.149	
	p	0.000	0.178	0.000	0.000	0.000	0.000	0.022	0.000	0.000	0.601	0.011	0.000	0.000	0.000	0.004	0.000	0.000	0.070	
	N	139	50	18	128	130	53	28	53	122	149	149	149	149	149	93	103	149	149	
PCB	**r**	0.037	-0.463**	-0.500*	0.226*	0.313**	0.676**	-0.222	0.804**	0.471**	0.817**	0.540**	0.661**	0.695**	0.455**	0.556**	0.168	0.646**	0.669**	0.085
	p	0.669	0.003	0.035	0.013	0.000	0.000	0.256	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.107	0.000	0.000	0.319
	N	139	40	18	120	120	53	28	53	112	139	139	139	139	139	93	93	139	139	140

Table 16. Linear regressions for methods against environmental and contaminant data. r: Pearson; p: significance (p-value); N: number of pairs of data included in the analysis. Significant correlations (p<0.05)are shaded in orange; highly significant regressions (p<0.01) are shaded in green; Very highly significant regressions (p<0.001) are shaded in blue; Non-significant regressions (p>0.05) are striped; regressions with significant slope, but opposite to expected, are shaded and striped.

		Salinity	O_2	%Mud (Φ<20 μm)	%Mud	%Organic Matter	Total Organic C	REDOX potential	Al	As	Cd	Cr	Cu	Hg	Ni	Mn	Fe	Pb	Zn	PAH	PCB
M-AMBI	**r²**	0.022	0.081	-0.047	0.060	0.004	0.151	-0.019	0.304	-0.006	0.085	-0.005	-0.005	0.018	-0.003	-0.007	-0.002	0.054	0.094	-0.005	-0.005
	p	0.038	0.004	0.629	0.001	0.192	0.001	0.487	0.000	0.852	0.000	0.895	0.959	0.031	0.553	0.575	0.430	0.001	0.000	0.656	0.564
	df	151	90	17	180	182	64	27	64	162	201	189	189	201	201	93	143	201	189	149	139
IQI	**r²**	0.288	0.029		0.055	0.137	0.420	-0.060	0.165	-0.018	0.153	0.029	0.105	0.030	0.066	0.014	0.011	0.004	0.265	0.004	0.042
	p	0.000	0.178		0.016	0.000	0.000	0.503	0.002	0.891	0.000	0.066	0.001	0.048	0.007	0.157	0.189	0.248	0.000	0.253	0.033
	df	97	30		88	88	49	9	49	57	96	84	84	96	96	71	71	96	84	84	84
BAT	**r²**	0.143	-0.033	-0.046	-0.012	0.164	0.180	0.114	0.221	0.302	-0.017	-0.018	-0.150	-0.016	-0.007	-0.018	-0.032	-0.017	-0.011	0.028	0.028
	p	0.002	0.442	0.620	0.584	0.001	0.006	0.093	0.002	0.001	0.983	0.846	0.629	0.761	0.457	0.504	0.787	0.898	0.527	0.120	0.117
	df	59	3	17	59	59	35	17	35	28	59	53	53	59	59	30	30	59	53	53	53
TAsBeM	**r²**	0.149	-0.019	-0.066	0.025	0.127	0.612	-0.047	0.073	-0.010	-0.006	0.019	0.037	-0.006	0.017	0.008	-0.002	0.017	0.002	0.011	-0.008
	p	0.000	0.856	0.623	0.047	0.000	0.000	0.895	0.044	0.969	0.603	0.070	0.018	0.673	0.075	0.205	0.357	0.077	0.260	0.135	0.798
	df	119	51	12	116	116	42	22	42	102	130	123	123	130	130	82	93	130	123	113	112
QSB	**r²**	-0.022	0.255		-0.011	-0.006				0.264	0.249	0.025	0.114	0.197	0.015	0.016	0.033	0.307	0.358	0.254	-0.045
	p	0.516	0.000		0.646	0.466				0.000	0.000	0.098	0.002	0.000	0.147	0.265	0.069	0.000	0.000	0.002	0.707
	df	26	66		74	76				70	76	70	70	76	76	20	70	76	70	30	20
BEQI2	**r²**	0.022	0.061	0.329	0.110	0.005	0.283	-0.019	0.534	-0.005	0.026	0.001	-0.003	0.002	0.004	0.009	-0.007	0.027	0.032	-0.005	-0.007
	p	0.039	0.010	0.008	0.000	0.168	0.000	0.491	0.000	0.714	0.012	0.282	0.500	0.245	0.178	0.174	0.784	0.011	0.008	0.606	0.953
	df	151	90	17	180	182	64	27	64	162	201	189	189	201	201	93	143	201	189	149	139

Regarding the correlations between environmental and contaminants data, from Table 15, it can be concluded that most of contaminants concentrations (including also some pressure-related parameters such as organic matter and total organic carbon content) are correlated. Moreover, most of them are also significantly correlated to the physical structure of the sediment (mud content), probably due to the association of contaminants and organic matter to fine fractions of the sediment. Conversely, salinity and O_2 concentration in bottom water (percent saturation) are not significantly correlated (or the correlation coefficient is quite low) to most of the remainder parameters.

Regarding the regression analyses between methods and physico-chemical parameters, from table 16, it can be summarized that the response of the methods to environmental parameters and contaminant concentrations is variable. E.g.: all methods respond to total organic carbon in sediment or aluminum concentration (there are no data to test this relationship for QSB) whereas none of them responds to redox potential or concentrations of chromium, manganese and iron; for the remainder of combinations, some methods significantly respond to physico-chemical parameters and some others do not, without clear patterns which could be useful to classify the methods into groups of similar responses. However, there is high variability in the quality of the response (in terms of r^2) of methods to physico-chemical parameters (r^2=0.004-0.612), with a low average value (r^2=0.142), which indicates high dispersion of EQR values.

The linear regression analyses were repeated at Type level (taking the abovementioned typology classifications from the previous phase of IC), but the results did not improve noticeably.

The reason for such a high dispersion of EQR values can be from differences in the resolution of data. I.e., for most of methods the EQR values were calculated at replicate level, with physico-chemical data at sample level. This led to relatively wide ranges of EQR values for single (or similar) values of environmental and contaminant data, which is tested further on.

STEP 3: Sediment chemical quality data: station/location level

In order to adjust the resolution of the data, the EQR values were averaged at station/location level (after confirming with the experts concerned that this was acceptable) and the linear regression analyses were repeated.

Unfortunately, the results did not improve a lot. In the analyses above, the relation between the benthic assessment approaches and Zn is most promising. Therefore, the regression analyses of this variable are looked at in more detail.

Table 17. Linear regressions for methods against environmental and contaminant data for each of TW Types defined in the previous phase of intercalibration. r: Pearson; p: significance (p-value); N: number of pairs of data included in the analysis. Significant correlations (p<0.05)are shaded in orange; highly significant regressions (p<0.01) are shaded in green; Very highly significant regressions (p<0.001) are shaded in blue; Non-significant regressions (p>0.05) are striped; regressions with significant slope, but opposite to expected, are shaded and striped.

			Salinity	O_2	%Mud (Φ<20 µm)	%Mud	%Organic Matter	Total Organic C	REDOX potential	Al	As	Cd	Cr	Cu	Hg	Ni	Mn	Fe	Pb	Zn	PAH	PCB
Type D	M-AMBI	r^2	0,083	0,163	-0,047	0,095	-0,012	0,340	-0,013	0,346	0,028	-0,011	-0,001	-0,012	-0,007	0,004	0,022	-0,020	-0,011	-0,001	0,003	0,007
		p	0,004	0,335	0,629	0,006	0,661	0,000	0,391	0,000	0,106	0,850	0,348	0,925	0,524	0,252	0,149	0,951	0,951	0,343	0,259	0,213
		df	87	3	17	68	68	44	17	44	58	88	83	83	88	88	50	50	88	83	83	83
	IQI	r^2	-0,018	-0,869		-0,017	-0,030	0,739		0,747	-0,070	0,235	0,224	0,111	0,133	0,132	0,239	0,027	0,186	0,268	0,043	0,188
		p	0,637	0,836		0,513	0,882	0,000		0,000	0,705	0,001	0,001	0,022	0,009	0,009	0,001	0,160	0,002	0,000	0,107	0,003
		df	43	2		34	34	29		29	13	43	38	38	43	43	38	38	43	38	38	38
	BAT	r^2	0,134	-0,033	-0,046	-0,016	0,140	0,193	0,114	0,260	0,302	-0,017	-0,018	-0,015	-0,014	-0,010	-0,018	-0,032	-0,016	-0,011	0,028	0,028
		p	0,003	0,442	0,620	0,788	0,002	0,005	0,093	0,001	0,001	0,888	0,846	0,629	0,674	0,527	0,504	0,787	0,780	0,527	0,120	0,117
		df	58	3	17	58	58	34	17	34	28	58	53	53	58	58	30	30	58	53	53	53
	TAsBeM	r^2	0,061	0,211	-0,066	-0,001	-0,020	0,611	-0,091	0,060	-0,022	-0,012	0,058	-0,002	-0,008	0,008	0,138	-0,005	-0,012	-0,010	-0,016	-0,012
		p	0,027	0,312	0,623	0,330	0,970	0,000	0,993	0,095	0,744	0,642	0,034	0,345	0,481	0,220	0,011	0,374	0,656	0,514	0,853	0,590
		df	64	3	12	50	50	31	12	31	41	65	60	60	65	65	39	39	65	60	60	60
	BEQI2	r^2	0,079	0,140	0,329	0,234	0,036	0,452	0,694	0,577	0,066	-0,008	0,043	-0,004	-0,011	0,052	0,117	-0,013	0,003	0,039	-0,012	-0,010
		p	0,005	0,347	0,008	0,000	0,064	0,000	0,000	0,000	0,028	0,601	0,032	0,419	0,979	0,018	0,008	0,561	0,270	0,040	0,833	0,712
		df	87	3	17	68	68	44	17	44	58	88	83	83	88	88	50	50	88	83	83	83
Type E	M-AMBI	r^2	0,113	0,164		0,015	-0,011	0,454		0,454	0,082	0,238	0,012	0,011	0,142	0,019	0,371	0,024	0,135	0,230	-0,013	-0,033
		p	0,024	0,000		0,135	0,874	0,014		0,014	0,005	0,000	0,163	0,174	0,000	0,105	0,001	0,094	0,000	0,000	0,484	0,848
		df	36	70		84	86	10		10	80	86	80	80	86	86	25	75	86	80	40	30
	IQI	r^2	0,068	-0,014		-0,017	-0,027	-0,041		-0,041	-0,044	-0,036	-0,045	-0,044	-0,032	-0,036	0,097	0,110	-0,017	-0,029	-0,025	-0,044
		p	0,088	0,384		0,478	0,622	0,458		0,458	0,873	0,950	0,910	0,883	0,759	0,995	0,105	0,091	0,477	0,558	0,514	0,869
		df	29	13		29	29	10		10	23	29	23	23	29	29	18	18	29	23	23	23
	TAsBeM	r^2	0,010	-0,025		0,016	-0,024				0,025	-0,025	0,114	0,095	0,001	0,103	-0,019	0,112	-0,013	-0,027	-0,035	-0,040
		p	0,268	0,616		0,213	0,739				0,170	0,787	0,022	0,033	0,315	0,026	0,470	0,025	0,479	0,838	0,760	0,915
		df	27	31		38	38				37	38	37	37	38	38	25	36	38	37	27	26
	QSB	r^2	-0,022	0,255		-0,011	-0,006				0,264	0,249	0,025	0,114	0,197	0,015	0,016	0,033	0,307	0,358	0,254	-0,045
		p	0,516	0,000		0,646	0,466				0,000	0,000	0,098	0,002	0,000	0,147	0,265	0,069	0,000	0,000	0,002	0,707
		df	26	66		74	76				70	76	70	70	76	76	20	70	76	70	30	20
	BEQI2	r^2	0,084	0,137		0,023	-0,010	0,390		0,390	0,084	0,139	0,049	0,037	0,088	0,049	0,281	0,068	0,105	0,144	-0,016	-0,028
		p	0,046	0,001		0,090	0,706	0,024		0,024	0,005	0,000	0,026	0,047	0,003	0,022	0,003	0,013	0,001	0,000	0,540	0,664
		df	36	70		84	86	10		10	80	86	80	80	86	86	25	75	86	80	40	30
Type F	M-AMBI	r^2	0,519	0,467		0,231	0,351	0,330	-0,051	0,330	0,045	0,136	0,100	-0,027	-0,001	0,009	0,696	0,602	0,359	0,383	0,051	0,020
		p	0,000	0,002		0,007	0,001	0,062	0,476	0,062	0,169	0,036	0,068	0,549	0,332	0,281	0,000	0,000	0,001	0,001	0,144	0,236
		df	26	15		26	26	8	9	8	22	25	24	24	25	25	16	16	25	24	24	24
	IQI	r^2	0,555	0,296		0,481	0,433	0,077	-0,060	0,077	-0,028	0,489	-0,049	0,018	0,032	-0,009	0,207	0,039	0,370	0,457	0,102	0,873
		p	0,000	0,026		0,000	0,000	0,237	0,503	0,237	0,496	0,000	0,877	0,254	0,204	0,382	0,058	0,240	0,001	0,000	0,080	0,000
		df	23	13		23	23	8	9	8	19	22	21	21	22	22	13	13	22	21	21	23
	TAsBeM	r^2	0,516	0,462		0,310	0,416	0,164	-0,055	0,164	0,004	0,184	0,050	-0,042	0,024	-0,020	0,684	0,591	0,402	0,444	0,026	0,029
		p	0,000	0,002		0,002	0,000	0,153	0,487	0,153	0,309	0,017	0,147	0,855	0,218	0,480	0,000	0,000	0,000	0,000	0,212	0,205
		df	26	15		26	26	8	9	8	22	25	24	24	25	25	16	16	25	24	24	24
	BEQI2	r^2	0,562	0,469		0,204	0,337	0,385	-0,034	0,385	0,076	0,104	0,137	-0,018	-0,008	0,031	0,729	0,656	0,349	0,365	0,046	0,031
		p	0,000	0,002		0,010	0,001	0,044	0,426	0,044	0,109	0,060	0,039	0,453	0,379	0,192	0,000	0,000	0,001	0,001	0,156	0,197
		df	26	15		26	26	10	9	8	22	25	24	24	25	25	16	16	25	24	24	24

Figure 35 shows that the regression lines does not improve when the EQR data per samples are averaged to location. The spreading of the EQR values (0.1-1.2) at low Zn values, indicates that other factors were causing those variations. When the analyses are zoomed in on the data per type (*cf* type D and E data; Figure 35), the relations were worse in some cases, showing opposite trends or no trend.

The relations between the benthic assessment methods and Zn were also biased by three extreme Zn values, which were excluded in Figure 36. There, it is shown that there are significant negative trends in EQR values against Zn for all methods, except for QSB (not significant trend) and BAT (positive trend, but not significant). But, again, at low Zn values the variation in EQR values is still too high to assign them as 'benchmark' sites.

In order to check whether such a variation at low levels of Zn (or other contaminants) could be explained by some of the remainder indicators or pressure, multiple regression analyses (by type) were carried out, using the stepwise selection method. As the full set of contaminant data was not available for all samples, only those variables with data for more than 75% of samples were tested for this exercise. Moreover, when any of the contaminants showed a trend which was contrary to the expected (higher EQR values with high concentrations of contaminants or low concentration of O_2) in the first step of the selection procedure, it was removed and the exercise was repeated.

For Type D, the results show that, using only the samples with the full set of tested variables, the only method responding to such variables was IQI (Table 18). The regression was negative and significant, but the number of samples included was reduced to 10. When all available Hg data are used for a single regression, that is not significant any more (r^2=0.318; d.f.=10; p=0.071).

In Type D samples, and for the samples with the full set of tested variables, BEQI2 and M-AMBI are related to Cd and Ni. Both Cd and Ni present negative coefficients for both regressions (Table 18). When all available data are used for single regression, the pattern is repeated with negative and significant regressions, except the one of AMBI against Ni, which is negative, but not significant (r^2=0.07; d.f.=50; p=0.054).

Finally, regarding Type F, all methods respond positively to O_2 concentration, when only the samples with the full set of contaminant data are included in the analyses (Table 18). The regressions are positive and significant. However, when all available O_2 data are also significant.

A comparison of regression lines was carried out using Type F samples, with the EQRs from different methods as dependent variable and O_2 saturation as independent variable (Figure 37). The model was significant (r^2=0.45; d.f.=61; p=0.000) and it was found that there are not significant differences neither among the slopes, nor among the intercepts (Table 19).

Hence, it can be concluded that all methods significantly respond to dissolved O_2 in Type F estuaries and that all of them respond in the same way (no differences in intercept and slope).

Table 18. Results from the multiple regression analyses carried out using the reduced dataset (with averaged EQR values). Only variables with data for more than 75% samples were tested. r2 refers to the value corrected for the degrees of freedom (d.f.).

Type	Tested Variables	Method	Introduced Variables	r^2	d.f.	*p*-value
D	OM, Cd, Cr, Cu, Hg, Ni, Pb, Zn, PAH, PCB	***BAT***	-	-	-	-
		BEQI2	-	-	-	-
		IQI	Hg	0.60	9	0.009
		M-AMBI	-	-	-	-
		TAsBeM	-	-	-	-
E	O_2, OM, As, Cd, Cr, Cu, Hg, Ni, Pb, Zn	***BEQI2***	Cd, Ni	0.27	45	0.000
		IQI*	-	-	-	-
		M-AMBI	Cd, Ni	0.29	45	0.000
		QSB	-	-	-	-
		TAsBeM	-	-	-	-
F	O_2, OM, As, Cd, Hg, Pb, Zn, PAH, PCB**	***BEQI2***	O_2	0.50	13	0.003
		IQI	O_2	0.31	11	0.035
		M-AMBI	O_2	0.50	13	0.003
		TAsBeM	O_2	0.49	13	0.003

*Error in the final model because of a singularity in the covariances-matrix.

****** As, Cr, Cu and Ni were removed from the analyses as they regressed positively against one (or some) ot the methods in the first step of the variable selection.

Table 19. Results from the comparison of regression lines for the EQR values obtained with the methods applied in Type F TW, against dissolved O_2 saturation in bottom.

	Sum of Squares	d.f.	Mean Square	F-ratio	*p*-value
O_2	1.663	1	1.663	46.99	0.000
Intercepts	0.241	3	0.080	2.27	0.091
Slopes	0.118	3	0.039	1.12	0.351
Model	2.022	7			

Figure 35. Regression lines from each of the assessment methos (y-axis) against Zinc concentration (X-axis) for all data (most left), averaged sample data (left), average type D data (right), average type E data (most right). All regressions are significant except those for: BAT and TAsBeM; BEQI and M-AMBI in type D; and IQI in type E.

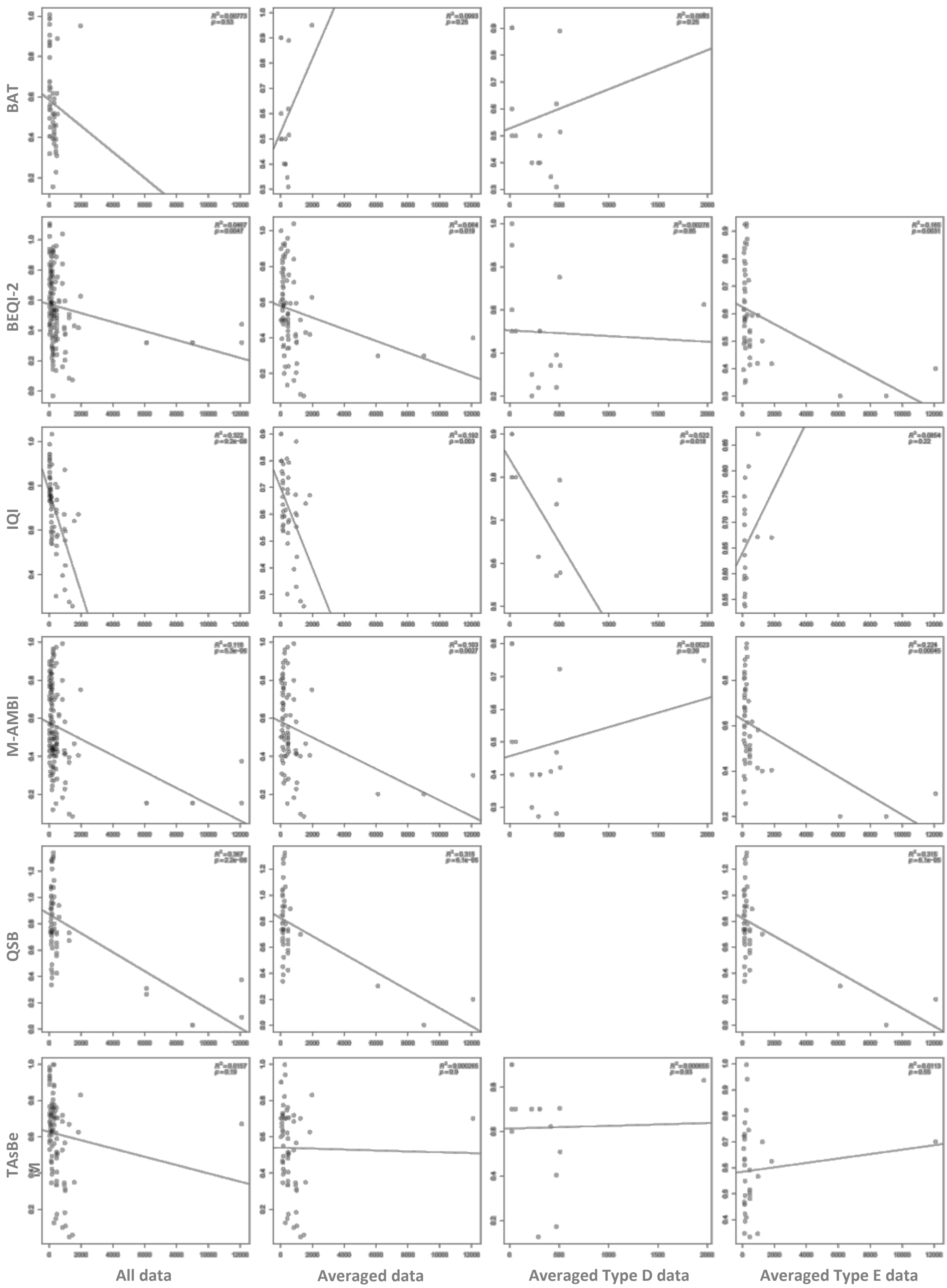


Figure 36. Scatter plots with linear trend lines for each of the assessment methos (y-axis) against Zinc concentration (X-axis) for all averaged sample data, excluding the data of the three highest Zinc concentration values.

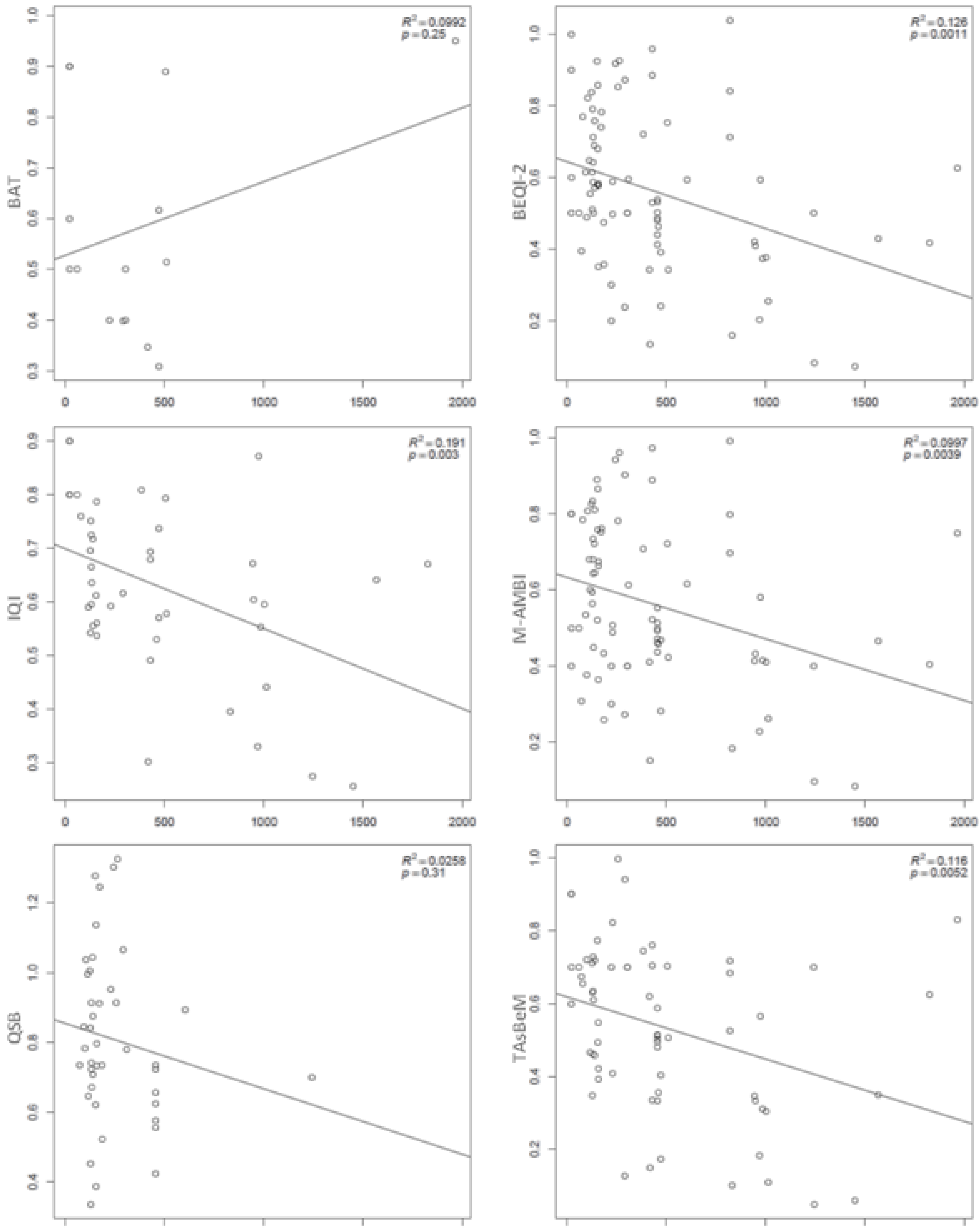
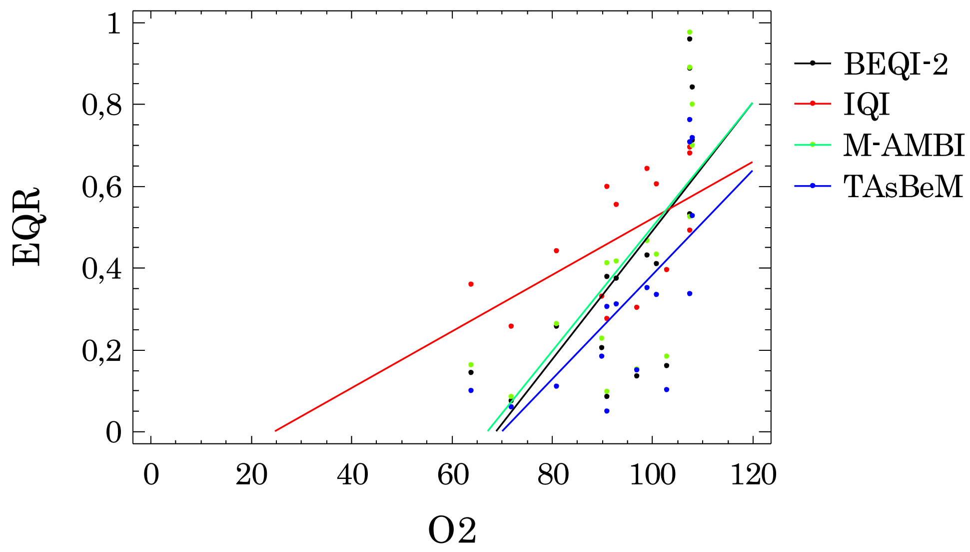


Figure 37. Regression lines of the EQR values obtained with each of the assessment methos against dissolved O2 concentration (X-axis) for the samples from Type F TW.



4.2.3 Conclusions

Intensive analyses on the entire and parts of the dataset, confirmed the ability of the benthic assessment methods to show pressure-response relations (similar for the methods in most cases) in certain cases. Nevertheless, those relations varied depending on the data availability and type, assessment method, pressure type and typology, which make it impossible to select 'common' benchmark sites, based on similar pressure levels across the transitional waters in the NEA-GIG region.

The reason for such a high dispersion of EQR values can be from differences in the resolution of data. I.e., for most of methods the EQR values were calculated at replicate level, with physico-chemical data at sample level. But the analyses on averaged data, show that this resolution in data plays only a minor role in the variation. The variation in EQR values against sediment chemical quality is mainly related to the multi-pressure environment of the transitional waters and the high natural variability of benthos in such systems. The regression of the methods against dissolved O_2 in Type F TW can be used as an example of the effect of the multi-pressure environment, as even the methods significantly respond to that, the intercept of the regression lines corresponds to negative values of EQR, with predicted EQR values close to 0 for O_2 saturation values around 25-70% (depending on the method used to derive the EQR).

The situation in the transitional waters of the NEA-GIG region seems to be unique in each waterbody, due to the fact that they are different in hydromorphological conditions and in the combination of pressures acting on the system status.

4.3 Assessment concept

All methods are based on similar parameters (diversity and presence/absence of sensitive and/or opportunistic species) and they are calculated at sample level (Table 20). Similar responses to pressures have been shown for all of them (as it has been demonstrated above), as they include equivalent metrics (AMBI or BO2A to account for opportunistic species; Margalef or species richness for abundance of species; and Shannon or Simpson evenness for diversity).

However, IQI attempts to account for within-habitat changes in metric values, where other MSs have fixed values for each habitat. In the case of the polyhaline where average salinity varies from 18 – 30 psu, reference conditions for the IQI change accordingly, but those of other MSs do not. This results in a divergence in results between the IQI and other methods which is dependent upon where along the e.g. 18 – 30 salinity gradient the samples are from (likely to more closely correspond at the 30 end). Given that reference conditions for the other MSs are calculated as the upper 95%ile within each habitat, the metric values used for reference conditions are likely to be from the samples close to salinity = ~30, and therefore EQRs at the 30 end will be higher than at the 18 end for a given pressure. Salinity effects on EQRs values and its consideration in the IC process have been analyzed in Annex I.

On the other hand, TASBEM considers the species abundance at Family level (not at species level as the rest of methods) providing in some cases different status evaluation compared with the rest of methods based on species level.

Therefore, Methods are based on similar parameters, but IQI and TASBEM have been excluded of the Ic process (explanation provided above).

Table 20. Summary of the assessment concept for each of the methods. RC: reference conditions.

Method	Assessment concept	Remarks
BEQI2	Multimetric including species richness, Shannon and AMBI	
BAT	Factor Analysis including Margalef index, Shannon and AMBI	Ecotope-specific RC; assessment first at the ecotope level, then area-weighted average water body assessment
TaSBeM	Multimetric including BO2A and Margalef index at Family level; for low-intertidal and subtidal soft-bottoms	Annelids should be identified at the lowest taxonomic resolution needed to classify them as opportunistic or not; Margalef calculated at Family level
M-AMBI	Factor Analysis including species richness, Shannon and AMBI	Ecotope-specific RC for high status; for soft-bottoms
QSB	Multimetric including species richness, Bray-Curtis similarity, abundance of opportunistic species and total abundance	It is the only method including similarity to a reference species composition; ecotope-specific RC
IQI	Multimetric including ratios for observed/reference values for species richness, Simpson evenness (1-λ’) and AMBI	For soft-bottoms; RC could be adapted to different habitats and/or sampling methods

5. Collection of IC dataset and benchmarking

5.1 Dataset description

The number of samples available in the full database is summarized in Table 21. It should be highlighted that the biological samples available (benthic organisms’ abundances), do not correspond exactly with the EQR data available. Moreover there are no quantitative physico-chemical data (e.g., salinity or sediment grain size) for all the samples. Finally, as it has been noted above, pressure information was not necessarily assessed at sample level.

Regarding acceptance criteria (Table 22), it should be noted that biological data were standardized in terms of sampling area and sieving mesh-size. This information was available in the database and the quality of sampling and analytical methodology was ensured by each member state.

Table 21. Overview of the number of sites/samples/data values

Member State	Number of sites or samples or data values		
	Biological data	Physico- chemical data	Pressure data
BE	397	Only for ecotope classification	Semiquantitative
FR	1187	Only for ecotope classification	Semiquantitative
D	500 (only EQR values)	Only for ecotope classification	Semiquantitative
NL	1527	Only for ecotope classification	Semiquantitative
PT	2618	Only for ecotope classification	Semiquantitative
SP-An	218	Only for ecotope classification	Semiquantitative
SP-BC	99	Only for ecotope classification	Semiquantitative
SP-C	90	Only for ecotope classification	Semiquantitative
RoI & UK	529	Only for ecotope classification	Semiquantitative

Table 22. Overview of the data acceptance criteria used for the data quality control.

Data acceptance criteria	Data acceptance checking
Data requirements (obligatory and optional)	Biological data were standardized after sampling area and sieving mesh-size
The sampling and analytical methodology	Information on sampling method, sampling area, sieving mesh-size and sampling season available. Quality ensured by each Member State.
Level of taxonomic precision required and taxa lists with codes	Organisms identified at species level whenever possible. Taxa standardized and validated after (1) European Register of Marine Species (ERMS), (2) World Register of Marine Species (WoRMS), and (3) Fauna Europaea.
The minimum number of sites / samples per intercalibration type	4128 samples available for Type D, 481 samples for Type E and 1971 for Type F
Sufficient covering of all relevant quality classes per type	Type D: in general, ≤10% of samples in Bad and High status (32% in High status according to IQI; 18% in High status according to TaSBeM); according to different methods, 67-94% of samples in Poor to Good Status Type E: <5% of samples in Bad status; according to different methods, 75-86% of samples in Poor to Good Status, except for QSB, for which only 42% of samples are in Poor to Good, with 54% of samples in High status (≤25% for the rest of methods) Type F: ≤5% of samples in Bad status; according to different methods, 66-93% of samples in Poor to Good Status; <15% of samples in High status, except for IQI (34% in High status)

5.2 Common benchmark: IC reference conditions or alternative benchmark

Common reference conditions were not defined, as there was not a sufficient number of samples in near-natural conditions in the database. Each method defined reference conditions which were applied to the full dabase. This was done following expert judgement and modelling (correlation between metrics and/or physico-chemical parameters), together with some statistical approaches (e.g., percentiles).

An alternative procedure for the selection of benchmark sites need to be used in this intercalibration, because the guidance principle cannot be fulfilled using this common dataset: *The benchmarking process must use harmonized criteria independent of national classifications (i.e., countries cannot simply nominate the sites they classify as high status as being their benchmark sites without further checking).* The analyses on the common dataset showed that it was impossible to select 'common' benchmarks sites, based on similar pressure levels across the transitional waters in the NEA-GIG region. This is related to the high variation in the pressure-response of the methods, which depend on the data availability, data type, assessment method, pressure type and typology.

In this sense, it was proposed to select benchmark sites in basis on the expert judgment, as was done for the NEA-GIG coastal water intercalibration for benthic invertebrates. It was expected that, based on the knowledge of the coastal areas and the stations included in the dataset, the experts could indicate the stations that were under minor pressures (or with more distance from the focus of main disturbances). Therefore, the experts would be able to indicate on basis of their opinion (and not based on the methods results), the stations with minor pressures.

However, they were not able to do so. With some exceptions with few samples in good status in Cantabria (SP) after expert judgment and a some samples in moderate status in the Basque Country (SP), the remainder of experts were not able to select benchmark sites from the dataset.

5.3 Benchmark standardisation

Since the use of reference benchmarking and alternative benchmarking was not possible, it was tried to apply continuous benchmarking. This alternative requires relevant pressure data being available; hence, it was not possible to apply such a procedure using the full database. In turn, a reduced database could be useful for this. However, where variability in the data is low, this may not be particularly problematic, but for the highly variable NEA11 data (in terms of the biological, environmental and contaminant data, along with some slight variations in methodology) large sample numbers are essential for the results to be robust.

In the case of types E and F, the standarization is not neccesary as the dataset used for intercalibrating the methods MAMBI and QSB (TYPE E) and MAMBI and BAT (type F) was from the same Member State, (Spanish data set for Type E and Portuguese data set for type F). In the case of type D, a kappa analyses was done, and the results indicated a high and significative value of agreement (>0.6) between the methods to be intercalibrated, leading us to conclude that the pressures are acting on the estuaries ia sama way, producing a similar deviation from the reference conditions and therefore in the same ecological status in different areas, not being neccesary corrections due to regional differences.

6. Comparison of methods and boundaries

6.1 IC option and common metrics

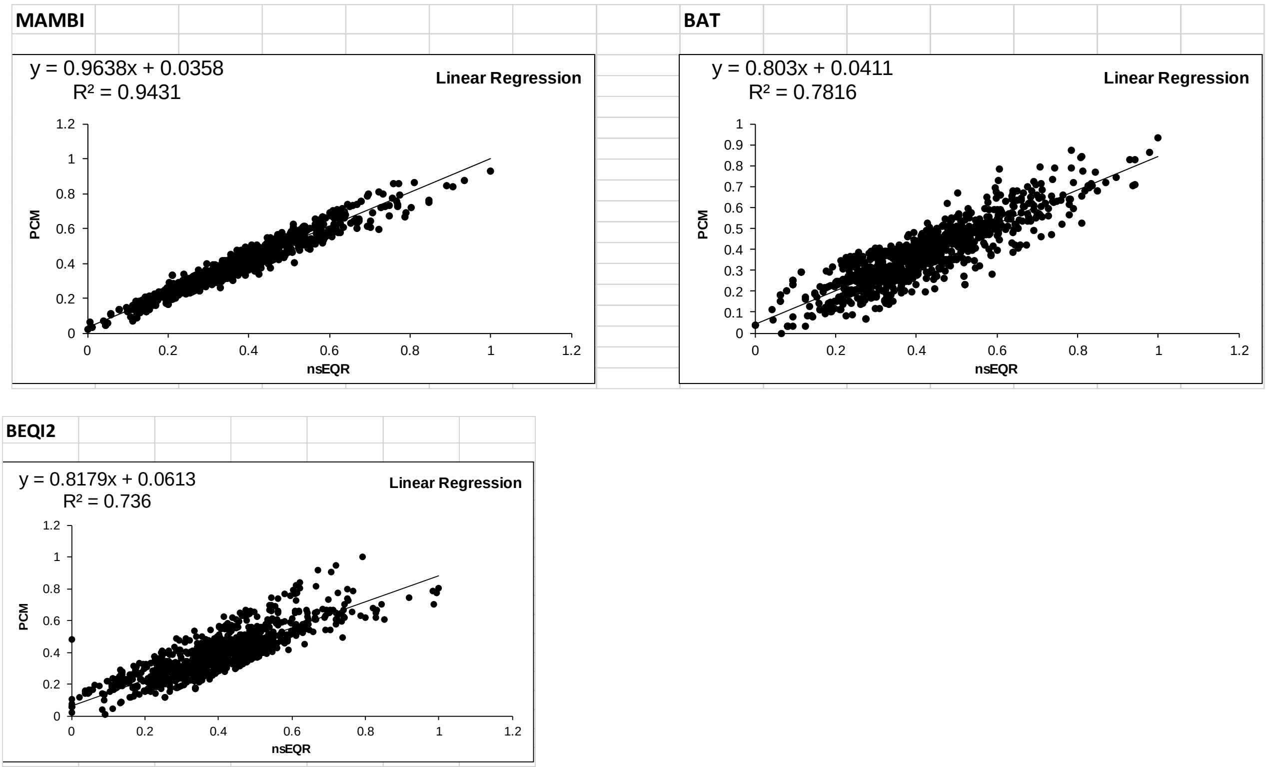
Option 3 was applied for all the types using the Intercalibration Excel Template Sheets for calculations (*IC_Opt3_sub_v1.24*, developed by Dirk Nemitz, Nigel Willby and Sebastian Birk, 2011). In the case of Types E and F was used the *Option 3-Two MS* Excel sheet for calculations. Previously, EQR values were normalized.

6.2 Results of the regression comparison

After including all data, regressions between each methodology and the pseudocommon metric were calculated.

Type D:

Figure 40. Regression results estimated for each assessment methods against the averaged values for the EQRs from the remainder of methods.

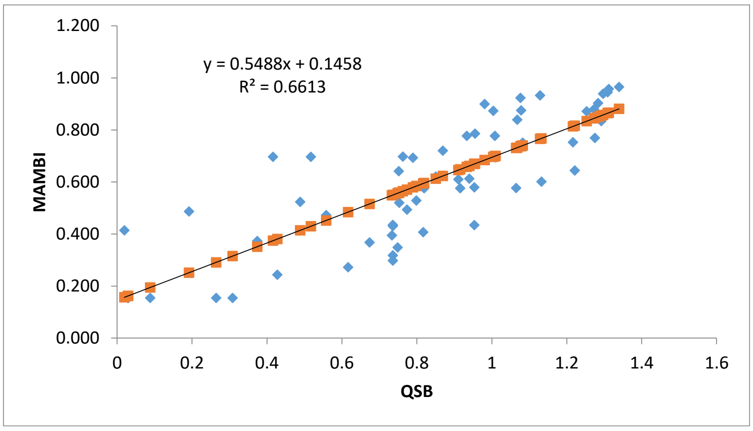


The relationships between the national methods and and the Psedo common metric are significant ($p<0.05$) and R2 are higher than 0.5.

Type E:

The relationship between M-AMBI and QSB is significant ($p<0.05$) and high ($R^2>0.5$) (Figure 38).

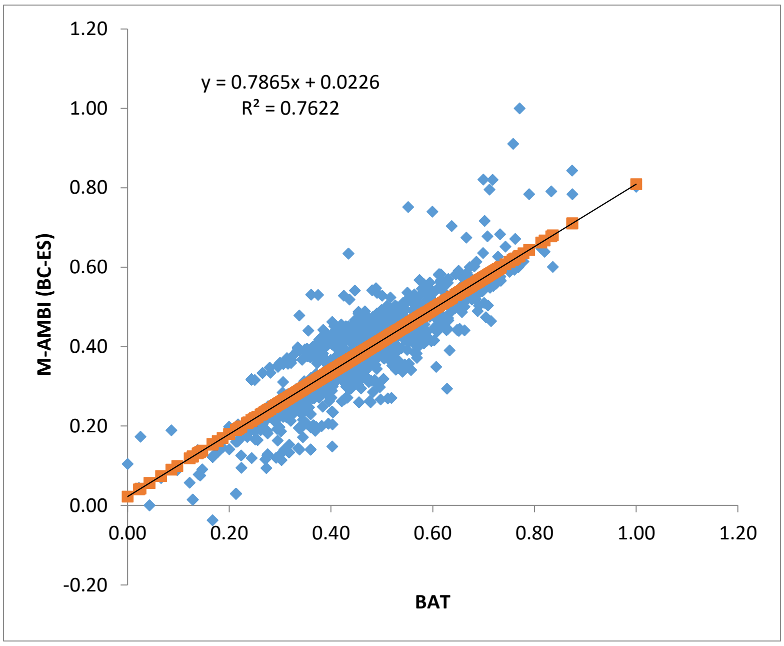
Figure 38. Regression results estimated between the assessment methods.



Type F:

The relationships between the national methods are significant ($p<0.05$). The R^2 value obtained are higher than 0.5 (Figure 39).

Figure 39. Regression results estimated between the assessment methods.



6.3 Comparability criteria

National boundaries exceeding a bias of 0.25 class equivalents should be adjusted to fall inside this allowed level of deviation, unless the provided thresholds are more stringent that the ones suggested after harmonization.

For **Type D**, M-AMB I(DE) and BAT (PT) the bias values exceed 0.25. However, the deviances for M-AMBI and are positive (Table 23). Hence, DE are not obliged to lower such boundaries. On the contrary, the deviances calculated for BAT are both negative. Thus, the H/G and G/M boundaries for BAT should be increased.

For **Type E** and Type F the biass values do not exceed 0.25 and adjustments are not necessary

Table 23. Values for the thresholds between quality classes for each of the intercalibrated methods and deviations from the thresholds suggested by the intercalibration common metric.Type D

	M-AMBI(DE)	**BAT**	**BEQI2**	**M-AMBI (SP)**
	Before adjustment			
Max	1.000	1.000	1.000	1.000
H/G	0.800	0.790	0.800	0.770
G/M	0.700	0.580	0.600	0.530
M/P	0.400	0.440	0.400	0.390
P/B	0.200	0.270	0.200	0.200
H/G bias	0.752	-0.420	-0.187	0.137
G/M bias	0.522	-0.312	-0.045	-0.055
	After adjustment			
H/G	0.800	0.838	0.800	0.850
G/M	0.600	0.582	0.600	0.700
H/G bias	0.752	-0.239	-0.187	0.137
G/M bias	0.522	-0.189	-0.045	-0.055

7. Final results to be included in the EC Decision

After the boundaries harmonisation, which is carried out using standardized EQR values, the results have to be reversed to the original EQR scale for each of the metrics. Hence, the opposite operation to the one used on the standardisation process has to be applied in order to re-establish the original range of values. In this sense, after that operation with offsets, the proposed H/G and G/M boundaries are the ones expressed on Table 24. These reults are included in the Part 1 of the EC Decision

Table 24. Values for the harmonized thresholds between High and Good (H/G), and Good and Moderate (G/M) quality classes for each of the intercalibrated methods and for each of the types.

Country	*National classification systems intercalibrated*	**Ecological Quality Ratios**					
		Type D		**Type E**		**Type F**	
		H/G	**G/M**	**H/G**	**G/M**	**H/G**	**G/M**
DE	***M-AMBI***	0.85	0.70	-	-	-	-
NL	***BEQI2***	0.80	0.60	-	-	-	-
PT	***BAT***	0.84	0.60	-	-	0.79	0.58
SP-BC	***M-AMBI***	-	-	0.77	0.53	0.77	0.53
SP-C	***QSB***	-	-	0.80	0.60	-	-

Boudaries values of the methods non intercalibrated due to justifieded reasons (see above) and accepted by the review panel are included in the Part 2 of the EC Decision

8. Correspondence common types versus national types

The results are directly applicable to the national types included in the IC common types.

9.Gaps of the current intercalibration

The Intercalibration has been not possible for the TAsBeM and IQI methods as already has been noted in previous sections.

France has not developed/proposed any method.

10.Ecological characteristics

Description of reference or alternative benchmark communities

The description of the benthic community characteristics at reference or alternative benchmark is summarized in Table 25. This information is generated from the WISER database.

Description of good status communities

The description of the benthic community characteristics at good status is summarized in Table 25. This information is generated from the WISER database.

Table 25. Overview of the description by the methods of the macro-invertebrate reference community and good status community

Method	*Description of reference community*	*Description of good status community*
BAT	Reference condition macrobenthic communities are dominated by pollution sensitive taxa (AMBI Ecological Group (EG) I taxa), have low relative abundance of indifferent (EG II) and tolerant (EG III) taxa and negligible relative abundance of opportunist (EG IV) and pollution indicator (EG V) taxa. High numbers of taxa with an even abundance distribution throughout the community is also indicative of reference conditions.	Community species richness (Margalef) and equitability (Shannon-Wiener) values are slightly reduced in comparison to values under reference conditions. While variable according to habitat, community composition (as assessed by AMBI) is slightly unbalanced. Community composition still dominated by EG I and II taxa. Slight reduction of sensitive taxa (EG I), and slight increase on tolerant taxa (EG III).
BEQI2	Benthic communities, species numbers, diversity typically for the habitat (sediment, salinity, exposure)- low number of opportunistic species.	High portion of sensitive taxa, complex communities, low number of opportunists, high species number and high diversity assemblages.
IQI	Reference condition macrobenthic communities are dominated by pollution sensitive taxa (AMBI Ecological Group (EG) I taxa), have low relative abundance of indifferent (EG II) and tolerant (EG III) taxa and negligible relative abundance of opportunist (EG IV) and pollution indicator (EG V) taxa. High numbers of taxa with an even abundance distribution throughout the community is also indicative of reference conditions.	Taxa number and Simpsons evenness are slightly reduced in comparison to values under reference conditions, while variables according to habitat (community abundance as assessed by AMBI) are slightly unbalanced: sensitive taxa (EG I) abundance may range from high sub-dominant to absent; indifferent taxa (EG II) are of low sub-dominant abundance; tolerant taxa (EG III) of dominant abundance; abundance of opportunistic (EG IV) and indicator taxa (EG V) may range from negligible or low to comparable abundance with indifferent taxa (EG II).
M-AMBI	See: Borja, A., F. Aguirrezabalaga, J. Martinez, J.C. Sola, L. Garciaarberas & J.M. Gorostiaga, 2003. Benthic communities, biogeography and resources management. In: Borja, A. & M. Collins, (Ed.). Ocenaography and Marine Environment of the Basque Country, Elsevier Oceanography Series n. 70: 27-50	See: Borja, A., A.B. Josefson, A. Miles, I. Muxika, F. Olsgard, G. Phillips, J.G. Rodríguez & B. Rygg, 2007. An approach to the intercalibration of benthic ecological status assessment in the North Atlantic ecoregion, according to the European Water Framework Directive. Marine Pollution Bulletin 55: 42-52.
QSB	-	-
TASBEM	-	-

11. Conclusions

Transitional water bodies were classified into 6 different types (A to F) according to size, river flow and intertidal area (relative to total area of the water body) and 3 of them were included in the IC: large estuaries (Type D), small-medium estuaries with >50% intertidal area (Type E) and small-medium estuaries with <50% intertidal area (Type F).

Due to the high variability in environmental parameters even between transitional waters of the same type, and to differences in the resolution of pressure and biological data in the database used for IC, the correlations between assessment methods and anthropogenic pressures were poor, even for correlations that were already demonstrated in previous literature. In order to reduce such variability, a reduced dataset including quantitative pressure data was built. Data were averaged at station level.

Hence, IC was feasible for: BEQI2, IQI and M-AMBI (DE) in Type D; for M-AMBI (SP-BC) and QSB in Type E; and for BAT, IQI and M-AMBI (SP-BC) in Type F. Since reference benchmarking and alternative benchamarking was nos possible, alternative benchmarking was applied. For Type D Cr concentration in sediment was used as common metric and Option 2 was applied for IC. For Type E, Pb concentration was selected as pressure indicator, and Option 3 was applied. Finally, for Type F, percent saturation of oxygen was used as pressure indicator and Option 3 was also applied.

After the appropriate options were applied: in Type D, IQI should adjust the good/moderate and the high/good boundaries (increase from 0.64 to 0.72 and from 0.75 to 0.83, respectively); in Type E, QSB should adjust the high/good boundary (increase from 0.80 to 0.85); and in Type F, BAT should adjust the good/moderate and the high/good boundaries (increase form 0.58 to 0.61 and from 0.79 to 0.81, respectively) and M-AMBI should adjust the good/moderate boundary (increase from 0.53 to 0.54). After such adjustments, all assessment approaches would meet the comparability criteria of the intercalibration guidance.

References

Aubry, A., M. Elliot, 2006. The use of environmental integrative indicators to assess seabed disturbance in estuaries and coasts: Application to the Humber Estuary, UK. *Marine Pollution Bulletin*, 53: 175-185.

Borja, A., B.G. Tunberg, 2011. Assessing benthic health in stressed subtropical estuaries, eastern Florida, USA using AMBI and M-AMBI. *Ecological Indicators*, 11: 295-303.

Borja, A.; Franco, F.; Valencia, V.; Bald, J.; Muxika, I.; Belzunce, M.J.; Solaun, O. 2004. Implementation of the European Water Framework Directive from the Basque Country (northern Spain): a methodological approach. *Marine Pollution Bulletin,* 48: 209-218.

Borja, A., I. Muxika, J.G. Rodríguez, 2009. Paradigmatic responses of marine benthic communities to different anthropogenic pressures, using M-AMBI, within the European Water Framework Directive. Marine Ecology, 30: 214-227.

Borja, A.; Barbone, E.; Basset, A.; Borgersen, G.; Brkljacic, M.; Elliott, M.; Garmendia, J.M.; Marques, J.C.; Mazik, K.; Muxika, I.; Neto, J.M.; Norling, K.; Rodríguez, J.G.; Rosati, I.; Rygg, B.; Teixeira, H.; Trayanova, A. 2011. Response of single benthic metrics and multi-metric methods to anthropogenic pressure gradients, in five distinct European coastal and transitional ecosystems. *Marine Pollution Bulletin*, 62: 499-513.

Callier, M.D., M. Richard, C.W. McKindsey, P. Archambault, G. Desrosiers, 2009. Responses of benthic macrofauna and biogeochemical fluxes to various levels of mussel biodeposition: An in situ "benthocosm" experiment. *Marine Pollution Bulletin*, 58: 1544-1553.

Clarke, K.R.; Warwick, R.M. 2001. Change in marine communities: an approach to statistical analysis and interpretation, 2nd edition. PRIMER-E: Plymouth.

Dauvin, J.C.; Ruellet, T. 2009. The estuarine quality paradox: is it possible to define an ecological quality status for specific modified and naturally stressed estuarine ecosystems? *Marine Pollution Bulletin*, 59: 38-47.

Hill, M.O. 1973. Reciprocal averaging: an eigenvector method of ordination. *Journal of Ecology*, 61: 237-249.

Hill, M.O.; Gauch, H.G. 1980. Detrended correspondence analysis: an improved ordination technique. *Vegetatio*, 42: 47-58.

Hill, M.O.; Šmilauer, P. 2005. *TWINSPAN for Windows version 2.3*. Centre for Ecology and Hydrology. University of South Bohemia, Huntingdon & Ceske Budejovice.

Muxika, I.; Borja, Á.; Bald, J. 2007. Using historical data, expert judgment and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Water Framework Directive. *Marine Pollution Bulletin*, 55: 16-29.

Muxika, I., P.J. Somerfield, Á. Borja, R.M. Warwick, 2012. Assessing proposed modifications to the AZTI marine biotic index (AMBI), using biomass and production. *Ecological Indicators*, 12: 96-104.

Neto, J. M., H. Teixeira, J. Patricio, A. Baeta, H. Veríssimo, R. Pinto, J. C. Marques, 2010. The response of estuarine macrobenthic communities to natural and human-induced changes: dynamics and ecological quality. Estuaries and Coasts, 33: 1327-1339

Phillips, G.R.; Anwar, A.; Brooks, L.; Martina, L.J.; Miles, A.C.; Prior, A. 2014. *Infaunal quality index: Water Framework Directive classification scheme for marine benthic invertebrates*. Report SC080016. UK Environment Agency: 193 pp.

Puente, A.; Juanes, J. A.; Echavarri-Erasun, B.; Galván, C.; Ondiviela, B. 2010. *A proposal for the assessment of the composition and community structure of benthic macroinvertebrates in transitional waters – The QSB index*. Abstract, ICES Congress, 2010: 7p.

Teixeira, H.; Neto, J.M.; Patrício, J.; Veríssimo, H.; Pinto, R.; Salas, F.; Marques, J.C. 2009. Quality assessment of benthic macroinvertebrates under the scope of WFD using BAT, the Benthic Assessment Tool. *Marine Pollution bulletin*, 58: 1477-1486.

Van den Bergh, E., S. Van Damme, J. Graveland, D.J. de Jong, I. Baten, P. Meire, 2003. *Studierapport natuurontwikkelingsmaatregelen ten behoeve van de Ontwikkelingsschets 2010 voor het Schelde-estuarium*, Hoofdstuk 2. Ecologische doelen voor het Schelde-estuarium, Werkdocument/RIKZ/OS/2003.825x, in Dutch.

Van Loon, W.M.G.M., A.R. Boon, A. Gittenberger, D.J.J. Walvoort, M. Lavaleye, G.C.A. Duineveld, A.J. Verschoor, 2015. Application of the Benthic Ecosystem Quality Index 2 to benthos in Dutch transitional and coastal waters. *Journal of Sea Research*, 103: 1-13.

Venice system (1959). *The final resolution of the symposium on the classification of brackish waters*. Archivio di Oceanografia e Limnologia, 11 (suppl): 243–248.

List of abbreviations and definitions

Key Terms:

Assessment method: The biological assessment for a specific biological quality element, applied as a classification tool, the results of which can be expressed as EQR.

Biological Quality Element (BQE): Particular characteristic group of animals or plants present in an aquatic ecosystem that is specifically listed in Annex V of the Water Framework Directive for the definition of the ecological status of a water body (for example phytoplankton or benthic invertebrate fauna).

Class boundary: The Ecological Quality Ratio value representing the threshold between two quality classes.

Common Intercalibration type: A type of surface water differentiated by geographical, geological, morphological factors (according to WFD Annex II) shared by at least two Member States in a GIG.

Compliance criteria: List of criteria evaluating whether assessment methods are meeting the requirements of the Water Framework Directive.

Ecological Quality Ratio (EQR): Calculated from the ratio observed value/reference value for a given body of surface water. The ratio shall be represented as a numerical value between zero and one, with high ecological status represented by values close to one and bad ecological status by values close to zero.

Geographic Intercalibration Group (GIG): Organizational unit for the intercalibration consisting of a group of Member States sharing a set of common intercalibration types.

Intercalibration: An exercise facilitated by the Commission to ensure that the high/good and good/moderate class boundaries are consistent with Annex V Section 1.2 of the Water Framework Directive and comparable between Member States.

IC Option: Option to intercalibrate (IC) different national assessment methods.

Method Acceptance Criteria: List of criteria evaluating whether assessment methods can be included in the intercalibration exercise.

Pressure: Human activities such as organic pollution, nutrient loading or hydromorphological modification that have the potential to have adverse effects on the water environment.

Reference/Benchmark sites: Reference sites meet international screening criteria for undisturbed conditions. Benchmark sites meet a similar (low) level of impairment associated with the least disturbed or best commonly available conditions.

Water Framework Directive: Directive 2000/60/EC establishing a framework for Community action in the field of water policy.

Abbreviations:

BE: Belgium

CA: Correspondence analyses

Cap Dre: Capital Dredging

DE: Germany

Dis Vol: Maintenance Disposal Volume

EG: Ecological group

EQR: Ecological Quality Ratio

FR: France

GIG: Geographic Intercalibration Group

Hyd Cha: Hydromorphological Changes

IC: Intercalibration

MS: Member State

PT: Portugal

RC: Reference conditions

RoI: Ireland

Sed Che: Sediment Chemical Quality

SP-A: Spain-Andalusia Region

Spain-BC: Spain-Basque country

SP: Spain-Cantabria region

Tot Pre: Total Pressure index

UK: United Kingdom

WFD: Water Framework Directive

List of figures

Figure 1 ordination plot obtained from the detrended correspondece analysis carried out once outliers were removed. 15

Figure 2. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances. 16

Figure 3. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances. 17

Figure 4. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances.. 18

Figure 5. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances. 19

Figure 6. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type D transitional waters. 20

Figure 7. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type E transitional waters. 21

Figure 8. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type F transitional waters. 22

Figure 9. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type D transitional waters. X1, X2 and X3 represent the first three axis of ordination.. 23

Figure 10. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type E transitional waters 24

Figure 11. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type F transitional waters. 25

Figure 12. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type D transitional waters. X1, X2 and X3 represent the first three axis of ordination. The minimum stress value for three dimensions was 0.330. Black: samples from intertidal sampling stations; Red: samples from subtidal sampling stations. 26

Figure 13. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type E transitional waters.. 27

Figure 14. Non-metric multidimensional scaling plots obtained from the Bray-Curtis distances calculated based of fourth root abundances, for type F transitional waters.. 28

Figure 15. Comparison of M-AMBI EQR values direct in a dumpsite for dredged material (Dumping), in an adjacent area (drift) and in a reference area in the river Weser.. 31

Figure 16. Correlation between dissolved oxygen (measured at Schaar van Ouden Doel) and benthos EQR (BEQI2) for the ecotope Mesohaline-

Intertidal. This analysis was initially made using 99/1 percentile reference values. 31
Figure 17. Correlation between DIN (measured at Schaar van Ouden Doel) and benthos EQR (BEQI2) for the ecotope mesohaline-intertidal. 32
Figure 18. Correlation between the percentage of high-dynamic litoral ecotope in the Westerschelde (mostly polyhaline ecotope;. 32
Figure 19. Linear regression between multipressure index and BAT (Borja et al., 2011). 33
Figure 20. EQR and EQS variation given by BAT method in the period from 1990 to 2006 at the Mondego estuary, and table including the events to which the changes in BAT should be related. 33
Figure 21. Scatterplot for TaSBeM and pressure indicators (sediment soluble phosphorous (P) and transitional-LUSI (tLUSI_d)) in transitional Atlantic waters of Andalusia, using monitoring data. 34
Figure 22. Scatterplot matrix for TaSBeM and transitional-LUSI (tLUSI) in transitional Atlantic waters of Andalusia using historical data. 34
Figure 23. Linear regression between pressure index (Borja et al., 2011) and M-AMBI. 35
Figure 24. Linear relationship between density of mussels and M-AMBI (Callier et al., 2011). 35
Figure 25. Table summarizing the response of M-AMBI to changes in pressures (discharge removal, dredging, sediment disposal, land reclamation and marina construction) (Borja et al., 2009). 36
Figure 26. Exponential regression between M-AMBI and oxygen saturation in samples from estuaries submitted to recovery schemes. 36
Figure 27. Spearman correlation between the QSB index and the index of global pressure. 37
Figure 28. Normalised IQI values versus summed pressure scores for Basque Coast with Pearson correlation (data from Borja et al., 2011). 37
Figure 29. Normalised IQI values versus summed pressure scores for Lesina Lagoon with Pearson correlation (data from Borja et al., 2011). 38
Figure 30. Normalised IQI values versus summed pressure scores for Mondego Estuary with Pearson correlation (data from Borja et al., 2011). 38
Figure 31. Normalised IQI values versus summed pressure scores for Oslofjord with Pearson correlation (data from Borja et al., 2011). 39
Figure 32. Normalised IQI values versus summed pressure scores for Varna Bay with Pearson correlation (data from Borja et al., 2011). 39
Figure 33. Regression lines from each of the assessed methods (Y-axis) against sediment chemical quality (X-axis). Colours indicate the country each of the samples come from. 42
Figure 34. Comparison of regression lines for each of the methods against sediment chemical quality using all the available data in the database. 42
Figure 35. Regression lines from each of the assessment methos (y-axis) against Zinc concentration (X-axis) for all data (most left), averaged sample data (left), average type D data (right), average type E data (most right). 51

Figure 36. Scatter plots with linear trend lines for each of the assessment methos (y-axis) against Zinc concentration (X-axis) for all averaged sample data, excluding the data of the three highest Zinc concentration values. 52

Figure 37. Regression lines of the EQR values obtained with each of the assessment methos against dissolved O2 concentration (X-axis) for the samples from Type F TW. 53

Figure 41. Regression results estimated between the assessment methods. 58

Figure 42. Regression results estimated between the assessment methods. 58

List of tables

Table 1 Overview of the national assessment methods 5

Table 2 Overview of the metrics included in the national assessment methods... 7

Table 3 Overview of the sampling and data processing of the national assessment methods included in the IC exercise 9

Table 4 Overview of the methodologies used to derive the reference conditions for the national assessment methods included in the IC exercise 10

Table 5 Reference conditions defined for each of the methods and types (see table 9)... 11

Table 6 Explanations for national boundary setting of the national methods included in the Ic exercise 12

Table 7 List of the WFD compliance criteria and the WFd compliance checking process and results of the national methods included in the IC exercisse.. 12

Table 8 Results obtained from the correspondence analysis carried out once outliers were removed 15

Table 9. Main common IC types identified in NEA transitional water bodies...... 16

Table 10. IC types for which assessment methods are appropriate. RC: reference conditions... 28

Table 11. Pressures addressed by the national methods included in the IC exercise and overview of the relationship between national methods and the pressures... 29

Table 12. Results from the regressions for methods against pressures. 40

Table 13. Results from the regressions for methods against sediment chemical quality (Sed Che). 41

Table 14. Number of pollutant data (and environmental data) provided by each of the Member States and total number. 43

Table 15. Correlations between environmental and contaminant data, from the reduced dataset. 44

Table 16. Linear regressions for methods against environmental and contaminant data.. 46

Table 17. Linear regressions for methods against environmental and contaminant data for each of TW Types defined in the previous phase of intercalibration.. 48

Table 18. Results from the multiple regression analyses carried out using the reduced dataset (with averaged EQR values). 50

Table 19. Results from the comparison of regression lines for the EQR values obtained with the methods applied in Type F TW, against dissolved O_2 saturation in bottom. 50

Table 20. Summary of the assessment concept for each of the methods. RC: reference conditions. 54

Table 21. Overview of the number of sites/samples/data values.................... 55

Table 22. Overview of the data acceptance criteria used for the data quality control. 55

Table 24. Values for the thresholds between quality classes for each of the intercalibrated methods and deviations from the thresholds suggested by the intercalibration common metric.Type D 59

Table 27. Values for the harmonized thresholds between High and Good (H/G), and Good and Moderate (G/M) quality classes for each of the intercalibrated methods and for each of the types. 60

Table 28. Overview of the description by the methods of the macro-invertebrate reference community and good status community 61

Table 29. Spearman correlation coefficients and associated p-values (in brackets) for the correlations between each of the methods and salinity, including all the transitional water body types defined in the previous intercalibration phase and including each of them separately (Type D, Type E and Type F).. 72

Table 30. Spearman correlation coefficients and associated p-values (in brackets) for the correlations between each of the methods and salinity, at saline stretch level (i.e. fresh water, oligohaline, mesohaline, polyhaline and euhaline) including all the transitional water body types defined in the previous intercalibration phase and including each of them separately (Type D, Type E and Type F). 73

ANNEX I

In order to check whether the EQR values are influenced by salinity, Spearman correlations were carried out both using the entire database and splitting it according to the abovementioned types (Table 26). Previously, 0, 2.5, 11.5, 24 and 32.5 salinities were removed, as they were fixed as the central value for each of the salinity stretches defined (i.e. fresh water, oligohaline, mesohaline, polyhaline and euhaline) and not real measures taken on field.

Table 26. Spearman correlation coefficients and associated p-values (in brackets) for the correlations between each of the methods and salinity, including all the transitional water body types defined in the previous intercalibration phase and including each of them separately (Type D, Type E and Type F).

Method	All TW types	Type D	Type E	Type F
BAT	0.06 (0.000)	0.04 (0.032)	0.14 (0.009)	0.24 (0.000)
BEQI2	0.09 (0.000)	0.10 (0.000)	0.11 (0.035)	0.18 (0.000)
IQI	-0.08 (0.001)	-0.03 (0.198)	-0.05 (0.435)	-0.21 (0.006)
M-AMBI	0.08 (0.000)	0.07 (0.000)	0.17 (0.001)	0.20 (0.000)
QSB	0.10 (0.174)	no data	0.10 (0.174)	no data
TAsBeM	-0.07 (0.000)	-0.08 (0.000)	-0.03 (0.655)	0.03 (0.536)

The results showed that correlation was poor or null (|r|<0.25) for all methods and water body types. So, there is not a salinity effect on the EQR values for any of the assessed methods at water body level.

The analyses were repeated at salinity stretch level for each of the water body types (Table 27). When data were segregated at salinity stretch level, significant moderate correlations were found in all stretches (except fresh water) and water body types for one or more assessment methods. With a significance of 0.01 (penalizing for multiple comparisons), there were significant and, at least, moderate correlations between salinity and:

- *BAT*: in polyhaline stretches in Type F transitional water bodies.
- *BEQI2*: in mesohaline stretches in Type D transitional water bodies; and polyhaline stretches in Type E and Type F transitional water bodies.
- *IQI*: in euhaline stretches in Type F transitional water bodies (no data for Type E).
- *M-AMBI*: in oligohaline stretches in Type D transitional water bodies; and polyhaline stretches in Type E and Type F transitional water bodies. However, it should be noted that in the oligohalines stretches from Type D water bodies the correlation between M-AMBI and salinity is negative and not positive, as it would be expected if richness and diversity are higher in the areas with higher salinity values.
- *QSB*: in polyhaline stretches in Type E transitional water bodies (no data for Type D and Type F).
- *TAsBeM*: in oligohaline stretches in Type D transitional water bodies (no data for Type E and Type F); mesohaline stretches in Type E and Type F transitional water bodies; and euhaline stretches in Type F transitional water bodies. However, it should be noted that in the

oligohalines stretches from Type D water bodies and in the mesohaline stretches from Type E water bodies the correlation between TAsBeM and salinity is negative and not positive, as it would be expected if richness and diversity are higher in the areas with higher salinity values.

Table 27. Spearman correlation coefficients and associated p-values (in brackets) for the correlations between each of the methods and salinity, at saline stretch level (i.e. fresh water, oligohaline, mesohaline, polyhaline and euhaline) including all the transitional water body types defined in the previous intercalibration phase and including each of them separately (Type D, Type E and Type F). In yellow: significant moderate (0.25<|r|<0.5) correlations with p<0.01. In orange: significant good (0.5<|r|<0.75) correlations with p<0.01.

FRESH WATER				
Method	**All TW types**	**Type D**	**Type E**	**Type F**
BAT	0.01 (0.955)	0.87 (0.221)	no data	-0.03 (0.891)
BEQI2	0.02 (0.929)	0.87 (0.221)	no data	-0.02 (0.916)
IQI	no data	no data	no data	no data
M-AMBI	0.05 (0.790)	0.87 (0.221)	no data	0.01 (0.979)
QSB	no data	no data	no data	no data
TAsBeM	no data	no data	no data	no data
OLIGOHALINE				
Method	**All TW types**	**Type D**	**Type E**	**Type F**
BAT	0.04 (0.455)	-0.10 (0.103)	0.22 (0.564)	0.07 (0.772)
BEQI2	-0.01 (0.883)	-0.13 (0.028)	0.02 (0.949)	-0.06 (0.792)
IQI	no data	no data	no data	no data
M-AMBI	-0.22 (0.000)	-0.34 (0.000)	0.25 (0.508)	0.06 (0.818)
QSB	-0.74 (0.050)	no data	-0.74 (0.050)	no data
TAsBeM	-0.36 (0.000)	-0.36 (0.000)	no data	no data
MESOHALINE				
Method	**All TW types**	**Type D**	**Type E**	**Type F**
BAT	0.12 (0.000)	0.20 (0.000)	-0.05 (0.600)	-0.07 (0.455)
BEQI2	0.09 (0.001)	0.25 (0.000)	-0.16 (0.101)	-0.01 (0.888)
IQI	-0.17 (0.706)	-0.17 (0.706)	no data	no data
M-AMBI	0.13 (0.000)	0.22 (0.000)	-0.20 (0.041)	0.08 (0.449)
QSB	-0.02 (0.922)	no data	-0.02 (0.922)	no data
TAsBeM	0.22 (0.000)	0.21 (0.000)	-0.39 (0.004)	0.31 (0.002)
POLYHALINE				
Method	**All TW types**	**Type D**	**Type E**	**Type F**
BAT	0.04 (0.067)	-0.02 (0.436)	0.23 (0.001)	0.41 (0.000)
BEQI2	0.06 (0.005)	0.01 (0.784)	0.31 (0.000)	0.34 (0.000)
IQI	-0.07 (0.005)	-0.02 (0.470)	-0.05 (0.435)	-0.11 (0.180)
M-AMBI	0.05 (0.019)	0.01 (0.645)	0.33 (0.000)	0.29 (0.000)
QSB	0.45 (0.000)	no data	0.45 (0.000)	no data
TAsBeM	-0.04 (0.100)	-0.05 (0.072)	-0.03 (0.745)	0.15 (0.036)

EUHALINE				
Method	**All TW types**	**Type D**	**Type E**	**Type F**
BAT	-0.11 (0.178)	0.14 (0.166)	0.02 (0.941)	0.43 (0.026)
BEQI2	-0.13 (0.134)	0.19 (0.062)	-0.09 (0.723)	0.45 (0.019)
IQI	0.30 (0.029)	-0.06 (0.754)	no data	0.56 (0.007)
M-AMBI	-0.08 (0.130)	-0.06 (0.290)	-0.40 (0.040)	0.40 (0.037)
QSB	-0.39 (0.049)	no data	-0.39 (0.049)	no data
TAsBeM	0.01 (0.825)	-0.01 (0.869)	no data	0.55 (0.005)

In view of such results, and in order to solve the problem of the salinity effect on some of the assessment methods, a possible solution would be to go on with the IC exercise at saline stretch level, separately for each of the water body types, first focusing on those water body types and saline stretches where none of the assessment methods correlates with salinity, i.e.: fresh water in all transitional water body types; oligohaline stretches in Type E and Type F transitional water bodies; polyhaline stretches in Type D transitional water bodies; and euhaline stretches in Type D and Type F transitional water bodies.

However, this exercise could be repeated using sediment grain size or mud content in sediment, and using also the tidal level, and more groups could be found. This would lead to an unmanageable number of groups and "intercalibrations". Moreover, some of the saline stretches from some of the water body types would be left out from the exercise.

Hence, the way forward was not clear from those analyses:

- In general, there are not significant relations between the benthic assessment methods and salinity. This allows to intercalibrate without considering sub-types, based on salinity.
- The analyses per type showed that in some cases sub-types need to be discriminated, but that would lead to an unmanageable amount of intercalibrations.

Therefore, salinity stretches will not be taken into account for next steps in the intercalibration process.

JK-NA-29644-EN-N

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/784291
ISBN 978-92-79-99321-3