Harmonization and Networking for contaminant assessment in the Ionian and Adriatic Seas

Deliverable T2.1.2 – Harmonized Quality Controlled dataset focused on contaminants

Work Package T2 - Data collection and definition of common data outputs focused on contamination

Authors:
M. Lipizer, M.E. Molina Jack – OGS

Data contributors:
M. Giani – OGS, G. Giorgi – ISPRA, S. Matijevic, D. Ivankovic - IOF, C. Zeri, S. Iona - HCMR, B. Cermelj – NIB, M. Fafandjel - RBI, D. Joksimovic, A. Castelli, M. Mitric – UoM-IMBK, M. Poje – ARSO, R. Bakiu - AUT

October 2019 - Revised March 2020

How to cite: M. Lipizer, M.E. Molina Jack, M. Giani – OGS, G. Giorgi – ISPRA, S. Matijevic, D. Ivankovic - IOF, C. Zeri, S. Iona - HCMR, B. Cermelj – NIB, M. Fafandjel - RBI, D. Joksimovic, A. Castelli, M. Mitric – UoM-IMBK, M. Poje – ARSO, R. Bakiu – AUT 2020 WPT2 HarmoNIA Deliverable T2.1.2 – Harmonized Quality Controlled dataset focused on contaminants. Doi: 10.6092/d7b07fe0-2830-4c60-a7d2-5db21f511c60

Index

1. Introduction

All data gathered within the project, as well as unrestricted data available in the ADRION region through EMODnet Chemistry data portal (https://www.emodnet-chemistry.eu/), have undergone a standardized and common Quality Control (QC) process as defined by Activity 1.1, Deliverable T1.1.3 "Methodological proposal for data Quality Control procedures". As a result, all data of contaminants in the Adriatic - Ionian Regional have been harmonized and quality controlled according to a common procedure. Data distribution (Fig. 1) and all related metadata are accessible on HarmoNIA data portal: http://harmonia.maris2.nl/search. Datasets are available from the portal according to access conditions defined by the consortium.

Fig. 1: Adriatic - Ionian Regional dataset of contaminants: data distribution.

Standardization and harmonization in data QC improves coherence among information available through different data centres and consolidates the network of institutions adopting a joint data management system. At the same time, data QC is a multi-step procedure which is constantly improved, learning from experience gained during the process of validation.

2. Dataset description

Total number of datasets: 5.666

Total number of datasets from HarmoNIA: 1.235

Dataset temporal extent: 1979 – 2018 (Fig. 2)

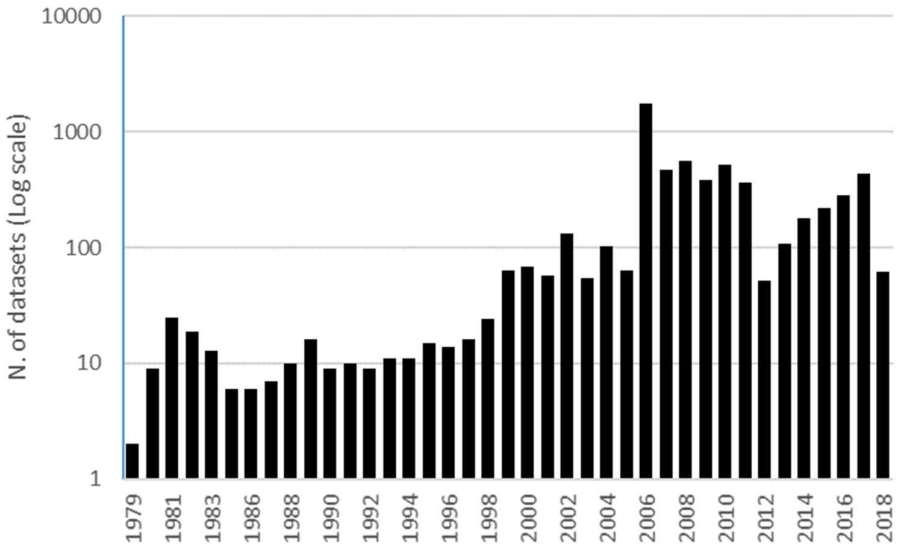


Fig. 2: Dataset temporal distribution (Logarithmic scale).

Data access restrictions:

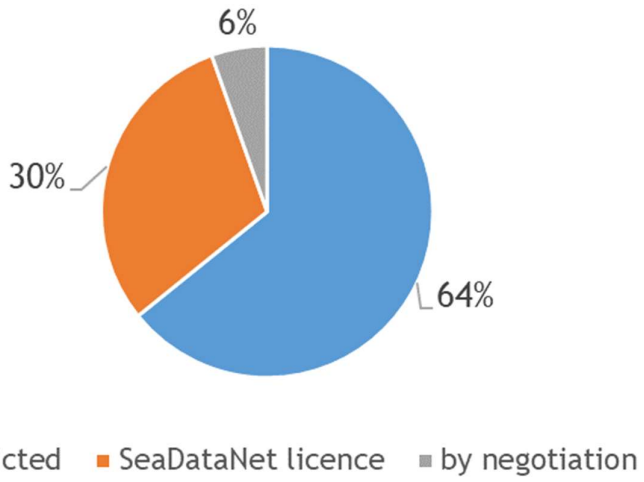


Fig. 3: Percentage of different conditions of data access (according to standard SeaDataNet access restriction policies, see Tab. 1 in the Annex 1).

Matrix categories:

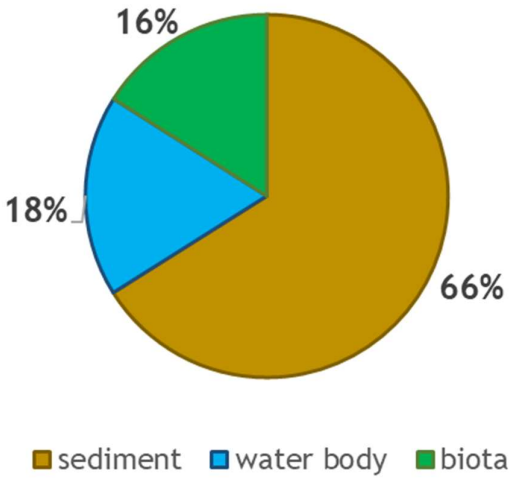


Fig. 4: Number of datasets per each matrix type.

Datasets per groups of variables (more details on the parameters available are reported in Tab. 2 in the Annex 1) (Fig. 5):

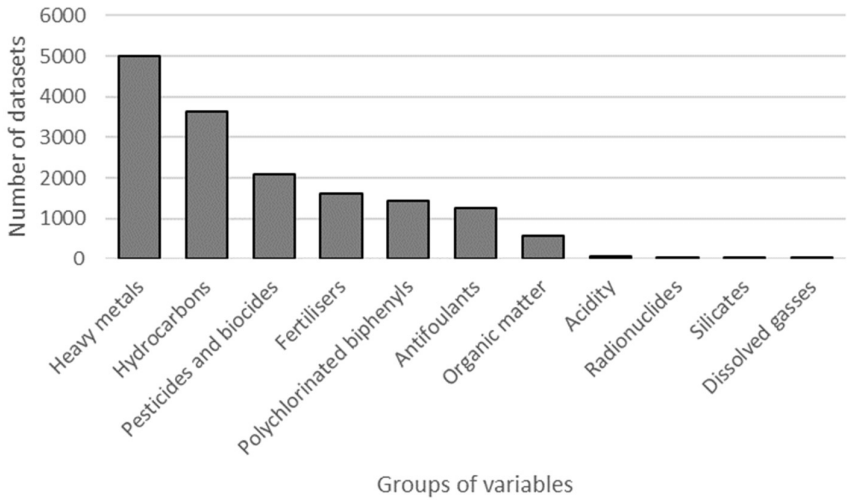


Fig. 5: Number of datasets per group of variables.

Dataset originators: 23 data originators from 8 countries (Table. 3 in Annex 1)

3. Data Quality Control process

Data Quality Control on the ADRION regional aggregated dataset involved a dataset formal control and a data quality control:

- **Formal control schema:**
  - ☐ ODV software used to perform all steps of format control in aggregated collections
  - ☐ Harmonization of the collection in terms of units and P01 parameters
  - ☐ Creation of "clean" collections in terms of stations and parameters in order to have 2 collections per matrix, one for time series and one for profiles. Parameters related to pore waters were discarded but another collection could have been created if considered relevant

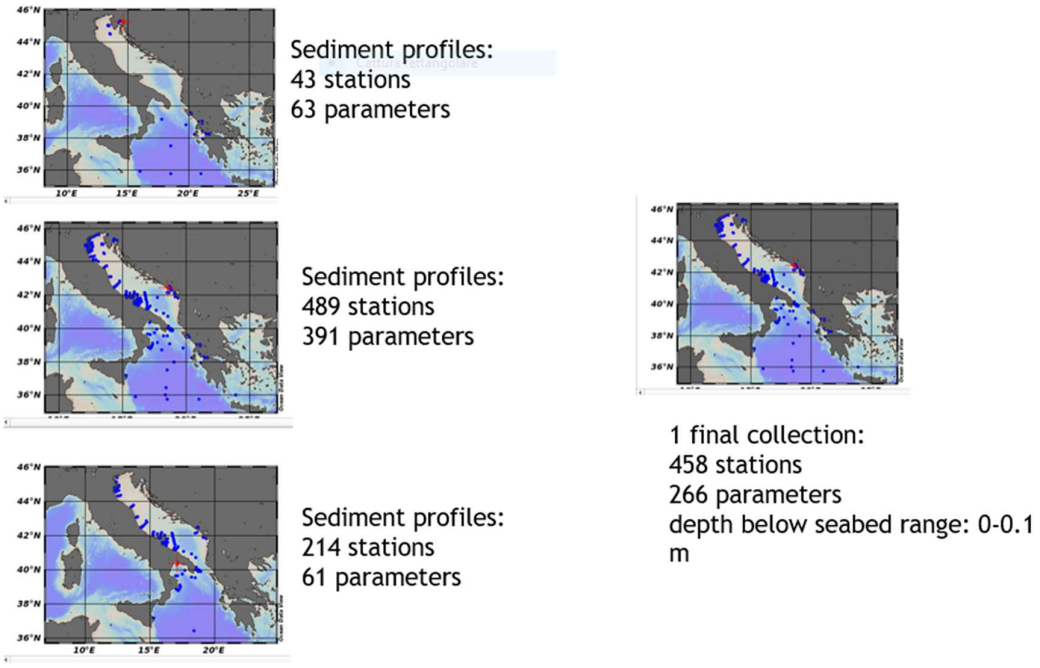


Fig. 6: Example for sediment datasets: creation of clean, coherent collections from the initial ones.

  - ☐ Deletion of variables with empty values for the whole collection
  - ☐ Check that 100% of stations have the correct primary variable. Contact originator if it is not present or it is wrong
  - ☐ Manual harmonization of variables in terms of units if needed
  - ☐ Manual aggregation of variables if needed

- ☐ Check values of quality flags:
  - Contaminant concentrations: value = 0 → QF=6
  - All parameters: empty value → QF=9
  - If conversions have been performed to achieve harmonized units: → QF=5
- ☐ Export filtered values for contaminants in sediment limited to measurements till 10 cm depth (coredist=0.1 m), considering that only that measurements are regarded for pollution assessment (discarding geological samples)

All the detailed steps of the formal control in HarmoNIA datasets are described in Annex 2

❖ **General Data quality control schema common to all matrices:**

- ☐ Use of ODV transposed and decomposed matrix (enables to filter per substance, matrix characteristics, analytical method,…)
- ☐ Relevant metadata availability check: eg. Time, depth,…
- ☐ Merging of "Time series" dataset with "Vertical profile" dataset to work on all data related to the same matrix
- ☐ Selection of substances (for which HarmoNIA has defined ranges, see Deliverable T1.1.3) to be QC
- ☐ Unit check: after harmonization STILL not uniform units
- ☐ If possible: manually harmonize units (eg. From mg/kg to µg/kg)
- ☐ Inspection of «0» values in concentrations and comparison with associated QF: this MUST be Q or 6
- ☐ Selection of each substance for QC

❖ **Data quality control schema for the sediment matrix:**

- ☐ Inspect matrix characteristics: eg. Total sediment, <2000 µm, < 63 µm (what to do?)
  - Apply ranges to the whole group of data
  - Apply ranges to a subgroup of data
- ☐ Plot total group to identify macroscopic discrepancies (Fig. 7)

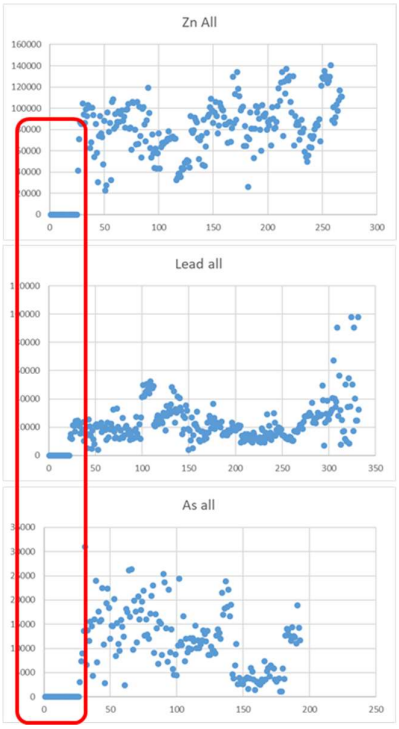


Fig. 7: Visual inspection and identification of anomalies, all related to the same data originator (possible un-correct data).

- ☐ Comparison with ranges (Deliverable T1.1.3)
- ☐ Assignment of Quality Values to data and contact with data originators for clarification and/or correction (Fig. 8).
- ☐ Due to the high number of pollution “hot spot” stations in the dataset and to the limited knowledge of contaminant concentrations in polluted areas, QC was focused on lower limits.

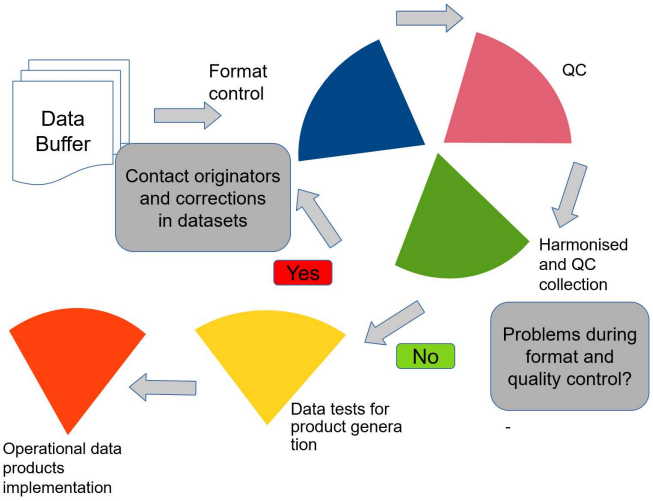


Fig. 8: Data Quality iterative process.

The Harmonized Quality Controlled dataset will be used for the generation of data products in different areas, chosen according to best temporal and spatial data coverage.

4. Data QC result

4.1. Sediment:

All HarmoNIA data and unrestricted data available for the ADRION region through EMODnet Chemistry related to the sediment matrix have been Quality Controlled using the harmonized approach.

The original "Time series" dataset collection contained 104.714 lines, with 93 different chemical substances; the original "Vertical profile" dataset collection contained 5.803 lines, with 104 different chemical substances[1]. After selecting the sub-set of substances for QC and merging the two types of datasets, the final dataset contained 63.121 lines, which represent the number of data of contaminant concentrations (of 27 different substances: 10 heavy metals and 17 hydrocarbons[2]) measured in sediment samples collected in the ADRION area. All data available through HarmoNIA

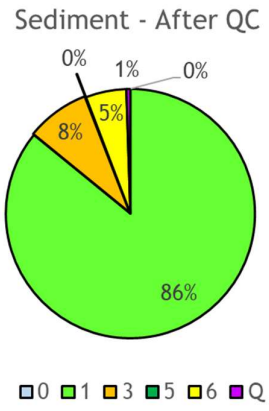
[1] The whole list of substances available in the dataset is presented in Tab. 4, Annex 1.
[2] The list of HarmoNIA Quality Controlled substances is listed in Tab. 5, Annex 1.

(http://harmonia.maris2.nl/search) are accompanied by a Quality value (QV[3]) indicated by the data providers. QV scale is based on SeaDataNet (Tab. 6 in Annex 1). After HarmoNIA Quality Control, several QV were modified. The results of HarmoNIA QC are summarized in Tab. 1, showing the original QV (first raw), QV assigned during HarmoNIA QC (second raw) and the difference (last raw). The overall quality of the dataset is shown in Fig. 9 and the data access conditions are shown in Fig. 10.

Data with no quality control (QV = 0) where assigned a QV = 1 (good data) if above the lowest limit (RMinC), QV = 6 (below detection limit) if equal to 0, QV = 3 (probably bad data) if lower than RMinC. In some cases, contact with data originators allowed to identify wrongly coded data, which were corrected, and these were flagged QV = 5 (modified data), according to SeaDataNet procedure.

Tab. 1: Number of data flagged with the different QV in the original dataset (upper row), after HarmoNIA Quality Control (second raw) and the difference (last raw). Total number of data: 63.121

	QV 0	**QV 1**	**QV 3**	**QV 5**	**QV 6**	**QV Q**
QV in original dataset	2.128	55.566	0	0	3.101	326
After HarmoNIA QC	0	54.246	5.193	21	3.335	326
difference	*2.128*	*3.320*	*-5.193*	*-21*	*-234*	*0*



[3] QV 0 = no quality control; QV 1 = good data; QV 3 = probably bad data; QV 5 = modified data; QV 6 = data below Limit of Detection; QV Q = data below Limit of Quantification

Fig. 9: Overall Data Quality of HarmoNIA Quality Controlled dataset of contaminants in the sediment (Codes of QV in the footnote).

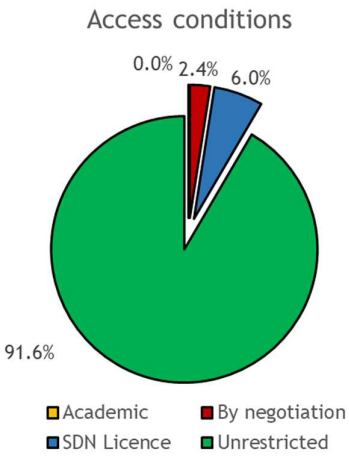


Fig. 10: Data access conditions of HarmoNIA Quality Controlled dataset of contaminants in the sediment.

4.2. Biota:

Due to unavailability of concentration ranges, data Quality Control of biota concerned:

- ☐ Inspection and, possibly, harmonization of measurement units,
- ☐ check of un-coherent measurement units,
- ☐ inspection of concentration values = 0 and flagging as QV=6

The whole list of substances present in HarmoNIA dataset of contaminant in biota is presented in Tab. 7 of Annex 1.

Tab. 2: Number of data flagged with the different QV in the original dataset (upper row), after HarmoNIA Quality Control (second raw). Total number of data: 10.312.

	QV 0	QV 1	QV 2	QV 3	QV 6	QV Q
QV in original dataset	980	7.800	5	1	1.295	230
After HarmoNIA QC	940	7.831	5	1	1.304	230

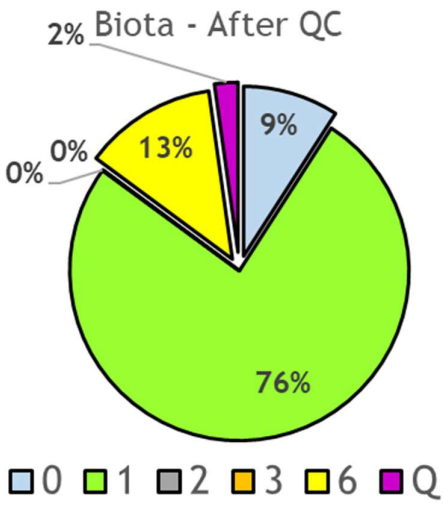


Fig. 11: Overall Data Quality of HarmoNIA Quality Controlled dataset of contaminants in biota (Codes of QV as before).

4.3. Seawater:

As for biota, due to unavailability of concentration ranges, data Quality Control of biota concerned:

- ☐ Inspection and, possibly, harmonization of measurement units,
- ☐ check of un-coherent measurement units,
- ☐ inspection of concentration values = 0 and flagging as QV=6

The whole list of substances present in HarmoNIA dataset of contaminants in seawater is presented in Tab. 8 of Annex 1.

Tab. 3: Number of data flagged with the different QV in the original dataset (upper row), after HarmoNIA Quality (second raw). Total number of data: 27.282.

	QV 0	QV 1	QV 6	QV Q
QV in original dataset	7.699	2.662	13.008	3.912
After HarmoNIA QC	6.353	1.944	15.072	3.912

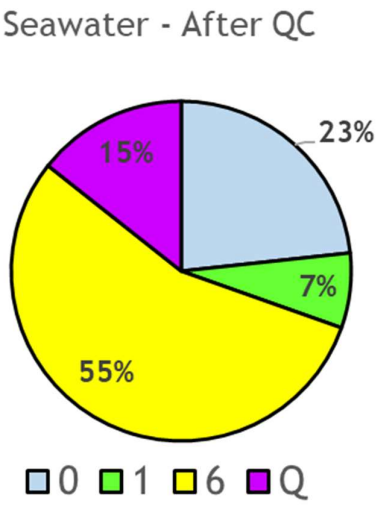


Fig. 12: Overall Data Quality of HarmoNIA Quality Controlled dataset of contaminants in seawater (Codes of QV as before).

In the case of data in biota and sediment, no sufficient knowledge is available to modify QV = 0 (meaning no quality control); as a consequence, 9% are still not quality controlled in the case of biota and 23% in the case of seawater.

As a remarkable result, Fig. 12 highlights that, in the case of seawater, 70% of data are below detection limit (55%) or below Limit of Quantification (15%).

Annex 1

Tab. 1: List of standard SeaDataNet data access restriction policies (Vocabulary L08: http://seadatanet.maris2.nl/v_bodc_vocab_v2/browse.asp?order=conceptid&formname=search&screen=0&lib=l08&v0_0=&v1_0=conceptid%2Cpreflabel%2Caltlabel%2Cdefinition%2Cmodified&v2_0=0&v0_1=&v1_1=conceptid&v2_1=3&v0_2=&v1_2=preflabel&v2_2=3&v0_3=&v1_3=altlabel&v2_3=3&v0_4=&v1_4=modified&v2_4=9&v0_5=&v1_5=modified&v2_5=10&x=33&y=16&v1_6=&v2_6=&v1_7=&v2_7=,

conceptid	preflabel	definition
CC	collection cost charge	A charge is made related to the cost of collecting the data.
FE	commercial charge	A charge significantly exceeding the cost of data collection and delivery is made for usage of the data.
LI	licence	Conditions of supply and usage of the data are specified in a formal agreement.
LS	SeaDataNet licence	Access to the data and usage are as specified in the SeaDataNet data policy and licence agreement
MO	moratorium	Data are initially restricted, but the access condition relaxes to academic or unrestricted once a specified period of time after an event (such as collection, publication, completion of QC procedures or project cessation) has elapsed.
NA	no access	Access to the data cannot be negotiated.
NC	distribution cost charge	A charge is made to cover the costs of delivering data to the user.
OG	organisation	The data are unrestricted to members of an organisation or a virtual organisation (such as project or cruise participants) but restricted to anybody else.
RS	by negotiation	The data are witheld from general circulation and disclosure but access may be obtained on a case-by-case basis through negotiation.
SR	academic	The data are freely available for research and education purposes. Usage acknowledgement is usually expected.
UK	unknown	The correct value is not known to and not computable by the creator of this information. However a correct value probably exists.
UN	unrestricted	The data are freely available to anybody and may be used for any purpose. Usage acknowledgement may be required.

Tab. 2: List of datasets per Discovery parameters (standard SeaDataNet P02 vocabulary, http://seadatanet.maris2.nl/v_bodc_vocab_v2/vocab_relations.asp?lib=P02)

Per Discovery Parameter	Datasets
Inorganic chemical composition of sediment or rocks	3216
Concentration of polycyclic aromatic hydrocarbons (PAHs) in sediment samples	2786
Metal concentrations in sediment pore waters	2322
Nitrogen concentrations in sediment	1600
Concentration of other organic contaminants in sediment samples	1588
Pesticide concentrations in sediment	1393
Concentration of polychlorobiphenyls (PCBs) in sediment samples	1247
Dissolved metal concentrations in the water column	867
Metal concentrations in biota	792
Organometallic and organometalloid species concentration parameters in sediments	785
Concentration of polycyclic aromatic hydrocarbons (PAHs) in the water column	526
Pesticide concentrations in water bodies	505
Concentration of other hydrocarbons in the water column	462
Concentration of other organic contaminants in the water column	448
Pesticide concentrations in biota	424
Organometallic and organometalloid species concentration parameters in water bodies	395
Concentration of organic matter in sediments	388
Concentration of polycyclic aromatic hydrocarbons (PAHs) in biota	226
Carbon concentrations in sediment	203
Total metal concentrations in water bodies	194
Concentration of polychlorobiphenyls (PCBs) in biota	156
Trace metalloid concentrations in biota	132
Concentration of other organic contaminants in biota	89
Organometallic species concentration parameters in biota	87
Carbonate chemistry in sediment pore waters	78
Geological sample radioactivity	35
Concentration of polychlorobiphenyls (PCBs) in the water column	34
Concentration of aliphatic hydrocarbons in sediment samples	28
Nutrient concentrations in sediment pore waters	13
Particulate metal concentrations in the water column	11
Alkalinity,	7
Ammonium and ammonia concentration parameters in water bodies	7
Dissolved oxygen parameters in the water column	7
Dissolved total or organic phosphorus concentration in the water column	7
Nitrate concentration parameters in the water column	7
Nitrite concentration parameters in the water column	7
Particulate total and organic nitrogen concentrations in the water column	7
Phosphate concentration parameters in the water column	7
Silicate concentration parameters in the water column	7
Radioactivity in water bodies	3
Concentration of silicon species in the water column	2

Tab. 3: List of data originators

Per Originator	Country	Edmo_code Originator	Datasets
ISPRA-Institute for Environmental Protection and Research	Italy	3009	3511
Environmental Agency of the Republic of Slovenia	Slovenia	1755	513
Institute of Oceanography and Fisheries	Croatia	700	457
Center for marine research - Rudjer Boskovic Institute	Croatia	702	170
National Institute of Biology - Marine Biology Station	Slovenia	1229	137
CNR Institute of Marine Science (ISMAR) - Bologna	Italy	145	131
Hellenic Centre for Marine Research Institute of Oceanography (HCMR/IO)	Greece	164	101
CNR Institute of Marine Science (ISMAR) (Lesina)	Italy	124	91
Adriatic LNG	Italy	4997	80
CNR Institute of Marine Science (ISMAR) - Ancona	Italy	144	78
CNR Istituto di Scienze Marine (Sezione di Venezia - ex IBM)	Italy	108	69
National Research Council of Italy - Institute for Marine and Coastal Environment (IAMC) (Sezione Capo Granitola)	Italy	4809	61
OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) Department of Biological Oceanography	Italy	2431	60
Odessa Branch of SOI (State Oceanographic Institute)	Ukraine	931	48
Institute of Marine Biology (IMBK)	Montenegro	2432	45
GEOMAR Helmholtz Centre for Ocean Research Kiel	Germany	2947	38
Briese Schiffahrt	Germany	1569	30
ARPA Friuli-Venezia Giulia - Alto Adriatico Observatory	Italy	1010	15
OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) Division of Oceanography	Italy	120	14
CNR Institute for the Marine and Coastal Environment (IAMC) - Napoli	Italy	252	11
Marine Hydrophysical Institute	Ukraine	727	3
CEREGE	France	560	2
Hellenic Centre for Marine Research (HCMR)	Greece	3051	1

Tab. 4: List of substances present in HarmoNIA dataset of contaminant in the sediment, in “Profiles” dataset type (left) and in “Time Series” dataset type (right).

Profiles	Time Series
2,2',3,3',4,4'-hexachlorobiphenyl	1-ethylnaphthalene
2,2',3,4,4',5,5'-heptachlorobiphenyl	1-methylnaphthalene
2,2',3,4,4',5'-hexachlorobiphenyl	1-methylpyrene
2,2',4,4',5,5'-hexachlorobiphenyl	2,2',3,3',4,4',5-heptachlorobiphenyl
2,2',4,5,5'-pentachlorobiphenyl	2,2',3,3',4,4'-hexachlorobiphenyl
2,2',5,5'-tetrachlorobiphenyl	2,2',3,4,4',5,5'-heptachlorobiphenyl
2,3,3',4,4',5-hexachlorobiphenyl	2,2',3,4,4',5',6-heptabromodiphenyl ether
2,3',4,4',5-pentachlorobiphenyl	2,2',3,4,4',5'-hexachlorobiphenyl
2,4,4'-trichlorobiphenyl	2,2',4,4',5,5'-hexabromodiphenyl ether
2,4',5-trichlorobiphenyl	2,2',4,4',5,5'-hexachlorobiphenyl
2,4'-dichlorodiphenyldichloroethylene	2,2',4,4',5,6'-hexabromodiphenyl ether
2,4'-dichlorodiphenyltrichloroethane	2,2',4,4',5-pentabromodiphenyl ether
4,4'-dichlorodiphenyldichloroethane	2,2',4,4',6-pentabromodiphenyl ether
4,4'-dichlorodiphenyldichloroethylene	2,2',4,4'-tetrabromodiphenyl ether
4,4'-dichlorodiphenyltrichloroethane	2,2',4,5,5'-pentachlorobiphenyl
4,4'-dichlorodiphenyltrichloroethane + 2,4'-dichlorodiphenyltrichloroethane	2,2',5,5'-tetrachlorobiphenyl
acenaphthene	2,3,3',4,4',5,5'-heptachlorobiphenyl
aldrin	2,3,3',4,4',5'-hexachlorobiphenyl
alpha-chlordane	2,3,3',4,4',5-hexachlorobiphenyl
alpha-hexachlorocyclohexane	2,3,3',4,4'-pentachlorobiphenyl
aluminium	2,3',4,4',5,5'-hexachlorobiphenyl
antimony	2,3,4,4',5-pentachlorobiphenyl
arsenic	2,3',4,4',5'-pentachlorobiphenyl
barium	2,3',4,4',5-pentachlorobiphenyl
benz(a)anthracene	2,3,6-trimethylnaphthalene
benzo(a)pyrene	2,4,4'-trichlorobiphenyl
benzo(b)fluoranthene	2,4'-dichlorodiphenyldichloroethane
benzo(e)pyrene	2,4'-dichlorodiphenyldichloroethylene
benzo(g,h,i)perylene	2,4'-dichlorodiphenyltrichloroethane
benzo(j)fluoranthene	2,6,10,14-tetramethylhexadecane
benzo(k)fluoranthene	2,6,10,14-tetramethylpentadecane
beta-hexachlorocyclohexane	2-methylnaphthalene
bismuth	3,6-dimethylphenanthrene
bromine	4,4'-dichlorodiphenyldichloroethane
C1-dibenzothiophenes	4,4'-dichlorodiphenyldichloroethylene

C1-phenanthrenes
C2-dibenzothiophenes
C2-phenanthrenes
C3-phenanthrenes
cadmium
caesium
cerium
chromium
chrysene + triphenylene
cobalt
copper
dibenzo(a,h)anthracene
dibenzothiophene
dieldrin
dysprosium
EPA 16 priority polycyclic aromatic hydrocarbons
erbium
europium
gadolinium
gallium
gamma-hexachlorocyclohexane
hafnium
hexachlorobenzene
holmium
inorganic carbon
ISMAR heptachlorobiphenyls
lanthanum
lead
lithium
lutetium
magnesium
molybdenum
naphthalene
neodymium
nickel
niobium
nitrogen
4,4'-dichlorodiphenyltrichloroethane
4,4'-dichlorodiphenyltrichloroethane + 2,4'-dichlorodiphenyltrichloroethane
acenaphthene
acenaphthylene
aldrin
alpha-hexachlorocyclohexane
aluminium
anthracene
arsenic
barium
benz(a)anthracene
benzo(a)pyrene
benzo(b)fluoranthene
benzo(e)pyrene
benzo(g,h,i)perylene
benzo(j)fluoranthene
benzo(k)fluoranthene
beta-hexachlorocyclohexane
cadmium
chromium
chrysene
copper
DDT+DDD+DDE
dibenzo(a,e)pyrene
dibenzo(a,h)anthracene
dibenzo(a,h)pyrene
dibenzo(a,i)pyrene
dibenzo(a,l)pyrene
dibutyltin
dieldrin
dioxin-like polychlorobiphenyls
endrin
EPA 16 priority polycyclic aromatic hydrocarbons
fluoranthene
fluorene
gamma-hexachlorocyclohexane
heptadecane

organic carbon
organic phosphorus
perylene
potassium
praseodymium
rubidium
samarium
scandium
silicon
silver
sodium
strontium
sulphur
terbium
thallium
thorium
thulium
tin
titanium
total iron
total manganese
total mercury
total phosphorus
tributyltin cation
tungsten
uranium
vanadium
water
ytterbium
yttrium
zinc
zirconium

hexachloro-1,3-butadiene
hexachlorobenzene
indeno(1,2,3-cd)pyrene
isodrin
lead
monobutyltin
naphthalene
nickel
octadecane
organic nitrogen
organic phosphorus
perylene
phenanthrene
pyrene
total iron
total manganese
total mercury
tributyltin cation
tributyltin compounds
vanadium
zinc

Tab. 5: List of HarmoNIA Quality Controlled substances in the sediment matrix.

Metals	As	**Hydrocarbons**	acenaphthene
	Cd		acenaphthylene
	Co		anthracene
	Cr		benz(a)anthracene
	Cu		benzo(a)pyrene
	Ni		benzo(b)fluoranthene
	Pb		benzo(g,h,i)perylene
	total_Hg		benzo(k)fluoranthene
	V		chrysene
	Zn		dibenzo(a,h)anthracene
			dibenzothiophene
			fluoranthene
			fluorene
			naphthalene
			perylene
			phenanthrene
			pyrene

Tab. 6: List of Quality Flags (QF) adopted according to SeaDataNet (L20 SEADATANET MEASURAND QUALIFIER FLAGS

QF Code	QF name	QF definition
0	no quality control	No quality control procedures have been applied to the data value. This is the initial status for all data values entering the working archive.
1	good value	Good quality data value that is not part of any identified malfunction and has been verified as consistent with real phenomena during the quality control process.
2	probably good value	Data value that is probably consistent with real phenomena but this is unconfirmed or data value forming part of a malfunction that is considered too small to affect the overall quality of the data object of which it is a part.
3	probably bad value	Data value recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.
4	bad value	An obviously erroneous data value.
5	changed value	Data value adjusted during quality control. Best practice strongly recommends that the value before the change be preserved in the data or its accompanying metadata.
6	value below detection	The level of the measured phenomenon was less than the limit of detection (LoD) for the method employed to measure it. The accompanying value is the detection limit for the technique or zero if that value is unknown.

7	value in excess	The level of the measured phenomenon was too large to be quantified by the technique employed to measure it. The accompanying value is the measurement limit for the technique.
8	interpolated value	This value has been derived by interpolation from other values in the data object.
9	missing value	The data value is missing. There should be no accompanying value in ODV format files. The accompanying value in SeaDataNet NetCDF data must be the absent data representation specified by the _FillValue parameter attribute and lie outside the range of data not flagged bad (4) or probably bad (3).
A	value phenomenon uncertain	There is uncertainty in the description of the measured phenomenon associated with the value such as chemical species or biological entity.
Q	value below limit of quantification	The level of the measured phenomenon was less than the limit of quantification (LoQ). The accompanying value is the limit of quantification for the analytical method or zero if that value is unknown.

Tab. 7: List of substances present in HarmoNIA dataset of contaminant in biota.

Time-Series

2,2',3,3',4,4',5,5'-octachlorobiphenyl
2,2',3,3',4,4',5-heptachlorobiphenyl
2,2',3,3',4,4'-hexachlorobiphenyl
2,2',3,4,4',5,5'-heptachlorobiphenyl
2,2',3,4,4',5',6-heptachlorobiphenyl
2,2',3,4,4',5'-hexachlorobiphenyl
2,2',4,4',5,5'-hexachlorobiphenyl
2,2',4,5,5'-pentachlorobiphenyl
2,2',5,5'-tetrachlorobiphenyl
2,3,3',4,4',5-hexachlorobiphenyl
2,3,3',4,4'-pentachlorobiphenyl
2,3',4,4',5-pentachlorobiphenyl
2,4,4'-trichlorobiphenyl
2,4'-dichlorodiphenyldichloroethane
2,4'-dichlorodiphenyldichloroethylene
2,4'-dichlorodiphenyltrichloroethane
4,4'-dichlorodiphenyldichloroethane
4,4'-dichlorodiphenyldichloroethylene
4,4'-dichlorodiphenyltrichloroethane
acenaphthene
acenaphthylene
alpha-hexachlorocyclohexane
anthracene
arsenic
benz(a)anthracene

Profiles

acenaphthene
anthracene
benz(a)anthracene
benzo(a)pyrene
cadmium
chromium
chrysene
copper
fluoranthene
indeno(1,2,3-cd)pyrene
lead
nickel
phenanthrene
pyrene
total mercury
zinc

benzo(a)pyrene
benzo(b)fluoranthene
benzo(e)pyrene
benzo(g,h,i)perylene
benzo(k)fluoranthene
cadmium
chromium
chrysene
copper
DDT+DDD+DDE
dibenzo(a,h)anthracene
dibutyltin
dieldrin
fluoranthene
fluorene
gamma-hexachlorocyclohexane
hexachloro-1,3-butadiene
hexachlorobenzene
indeno(1,2,3-cd)pyrene
lead
lipids
monobutyltin
naphthalene
nickel
phenanthrene
pyrene
selenium
silver
total iron
total manganese
total mercury
tributyltin cation
zinc

Tab. 8: List of substances present in HarmoNIA dataset of contaminant in seawater.

Profiles:	**TimeSeries**
1,1,2-trichloroethene	1,1,2-trichloroethene
1,2,3-trichlorobenzene	1,2,3-trichlorobenzene
1,2,4-trichlorobenzene	1,2,4-trichlorobenzene
1,2-dichloroethane	1,2-dichloroethane
1,3,5-trichlorobenzene	1,2-dimethylbenzene

1,3-dimethylbenzene and 1,4-dimethylbenzene	1,3,5-trichlorobenzene
2,3',4,4',5-pentachlorobiphenyl	1,3-dimethylbenzene and 1,4-dimethylbenzene
2,4-dichlorophenol	1-methylnaphthalene
3,3',4,4'-tetrachlorobiphenyl	1-methylphenanthrene
4,4'-dichlorodiphenyltrichloroethane	2,2',4,4',5,5'-hexabromodiphenyl ether
4-chloro-3-methylphenol	2,2',4,4',5,6'-hexabromodiphenyl ether
acenaphthylene	2,2',4,4',5-pentabromodiphenyl ether
alachlor	2,2',4,4',6-pentabromodiphenyl ether
aldrin	2,2',4,4'-tetrabromodiphenyl ether
aluminium	2,4,4'-tribromodiphenyl ether
anthracene	2,4'-dichlorodiphenyldichloroethane
arsenic	2,4'-dichlorodiphenyltrichloroethane
atrazine	4,4'-dichlorodiphenyldichloroethane
benz(a)anthracene	4,4'-dichlorodiphenyldichloroethylene
benzene	4,4'-dichlorodiphenyltrichloroethane
benzo(a)pyrene	alachlor
benzo(b)fluoranthene	aldrin
benzo(g,h,i)perylene	alpha-endosulfan
benzo(k)fluoranthene	alpha-hexachlorocyclohexane
bis(2-ethylhexyl)phthalate	aluminium
bromodichloromethane	ammonium
cadmium	anthracene
caesium-137	arsenic
chlorfenvinphos	atrazine
chlorpyrifos	barium
chromium	benz(a)anthracene
chrysene	benzene
cobalt	benzo(a)pyrene
copper	benzo(b)fluoranthene
dibenzo(a,h)anthracene	benzo(g,h,i)perylene
dibromochloromethane	benzo(k)fluoranthene
dichloromethane	bis(2-ethylhexyl)phthalate
dieldrin	cadmium
endrin	chlorfenvinphos
ethylbenzene	chloroalkanes C10-13
fluoranthene	chlorobenzene
fluorene	chlorpyrifos
hexachlorobenzene	chromium
indeno(1,2,3-cd)pyrene	chrysene
isodrin	copper
lead	DDT+DDD+DDE
methylbenzene	dibenzo(a,h)anthracene

molybdenum
naphthalene
nickel
pentachlorobenzene
pentachlorophenol
phenanthrene
pyrene
silver
simazine
styrene
tetrachloroethene
tetrachloromethane
total iron
total manganese
total mercury
total petroleum hydrocarbons
trichloromethane
trifluralin
vanadium
zinc

dibutyltin
dichloromethane
dieldrin
diuron
endrin
ethylbenzene
fluoranthene
fluorene
gamma-hexachlorocyclohexane
hexachloro-1,3-butadiene
hexachlorobenzene
indeno(1,2,3-cd)pyrene
isodrin
isoproturon
lead
methylbenzene
molybdenum
naphthalene
nickel
oxygen
pentachlorobenzene
pentachlorophenol
perylene
phenanthrene
phosphate
reactive mercury
silver
simazine
tetrachloroethene
tetrachloromethane
total iron
total mercury
tributyltin cation
tributyltin compounds
trichlorobenzenes
trichloromethane
trifluralin
xylene
zinc

Annex 2

During the import phase of HarmoNIA data buffer (CDI+ODV files), performed with ODV software, 6 collections were produced:

- 1 timeseries (with parameters for water, sediment and biota).
- 3 sediment profiles: primary variables COREDIST (Depth (spatial coordinate) relative to bed surface in the bed) (units P06 ULAA "Metres", ULCM "Centimetres", and ULCM "Cm").
- 2 ocean profiles: primary variables DEPHPR01 (Depth (spatial coordinate) relative to water surface in the water body by profiling pressure sensor and conversion to seawater depth using UNESCO algorithm) and ADEPZZ01 (Depth (spatial coordinate) relative to water surface in the water body).

For sediment profiles, it is necessary to harmonise units for depth.

For ocean profiles, DEPHPR01 and ADEPZZ01 can be considered equivalent, so DEPHPR01 can be edited into ADEPZZ01 for the whole collection.

For ocean profiles and sediment profiles, it is possible to import one collection into the other. A previous step is needed: It is necessary to add the data variables from the second collection to the main collection. This can be done through the properties of the collection→ data variables → new→ add from another collection. Then you start with the import dialog in the option "add/replace station data". For water, depth variables can be associated without any transformation (as there are in the same units and considered equivalent parameters). For sediment, it is necessary to convert the primary variable during the import to have the same units in the three collections.

Ocean profiles have 100% of Depth as primary variable.

Sediment profiles have 100% of COREDIST as primary variable.

Inside Ocean profiles there are 41 biota stations, that have to be selected and exported into a new collection, using the filter for the metadata field "variables measured". Export stations related to biota. Open each of the new created collections, delete the empty variables and export the collection using the option to obtain a SDN harmonized ODV collection.

For timeseries, parameters related to the three matrices must be splitted, using the filter for the metadata field "variables measured". Export stations related to biota,

sediment and water. Open each of the new created collections, delete the empty variables and export the collection using the option to obtain a SDN harmonized ODV collection. Some parameters of the collection may belong to other matrix if the station samples multiple matrices.

The 3 timeseries contain 100% of the primary variable.

For sediment timeseries, the collection contains additional parameters related to depth:

- Sea-floor depth in m (BATHDPTH: Sea-floor depth (below mean sea level) {bathymetric depth}; ULAA)
- Depth below seabed in m (COREDIST: Depth (spatial coordinate) relative to bed surface in the bed; ULAA)
- Depth below seabed in cm (COREDIST: Depth (spatial coordinate) relative to bed surface in the bed; ULCM)

For COREDIST, parameters with different units can be aggregated into one variable through view → derived variables → special → Aggregated variable → Add. Data variables to be aggregated are chosen and through the conversion option metres are conversed into centimetres through the general linear transformation, multiplying by 100. After that a new collection with the new created data variable and without the two original ones can be created (the collection contains only 25% of COREDIST data).

For water timeseries, the collection contains additional parameters related to depth:

- Sea-floor depth in m (BATHDPTH: Sea-floor depth (below mean sea level) {bathymetric depth}; ULAA): 13%
- Depth in m (ADEPZZ01: Depth (spatial coordinate) relative to water surface in the water body; ULAA): 25%

For biota timeseries, the collection contains additional parameters related to depth:

- Sea-floor depth in m (BATHDPTH: Sea-floor depth (below mean sea level) {bathymetric depth}; ULAA): 4%
- Depth in m (ADEPZZ01: Depth (spatial coordinate) relative to water surface in the water body; ULAA): 68%

- Depth in m (DEPHPR01: Depth (spatial coordinate) relative to water surface in the water body by profiling pressure sensor and conversion to seawater depth using UNESCO algorithm): 1 %
- MINWDIST in m (MINWDIST: Depth (spatial coordinate) minimum relative to water surface in the water body; ULAA): 1%
- MAXWDIST in m (MAXWDIST: Depth (spatial coordinate) maximum relative to water surface in the water body): 1%

Some parameters of the collections are in moles. It is necessary to prepare a list of molecular weight of the substances to perform the transformations. Through ODV → view → Derived variables it is possible to choose inside special transformations, the option converted variable and select the parameter and perform a general linear transformation with the molecular weight in the suitable units to obtain the desired ones. Once this is done, it is necessary to edit the variable in the collection and insert in the comment "Codes: SDN:P01::ADEPZZ01 SDN:P06::ULAA" with the suitable P01 and P06 codes. Then the export of the harmonized SDN collection can be performed.

Data variables measured in mg/kg for water won' t be transformed in mg/l as the conversion needs to use density and therefore this may introduce an error.

Inside the collections: water_TS, sediment_TS, biota_TS and sediment_PR, there are datasets with Bot.depth=0. SDN guidelines have been recently corrected in this sense and bot. depth must be left to empty if unknown or inapplicable instead of set to zero.

For contaminants variables, values=0 must be flagged as 6 or Q.