NCEA | Natural Capital and Ecosystem Assessment

PelCap - mNCEA NC34 Pelagic Programme

Year 2 (2023/24) Monitoring, sample analysis, data storage and data analysis

Date: March 2024

Authors: Mike Best (Environment Agency), Michelle Devlin (Cefas), David Johns (Marine Biological Association)

Environment Agency | Cefas | Marine Biological Association | JNCC Joint Nature Conservation Committee | Marine Management Organisation | Department for Environment Food & Rural Affairs

Project title:	NC34 Pelagic Programme; plankton monitoring (Year 2), evidence optimisation, and an investigation of the link between our plankton and fisheries natural capital
Lead organisation	Environment Agency
Project SRO/lead:	Roger Proudfoot
Project start date:	Y2 (of 3): 01 April 2023
Project end date:	Y2 (of 3): 31 March 2024

Project details:

Version history:

Version	Date	Summary
1.0	25/01/2024	First draft
1.5	18/03/2024	Updates for Cefas, PML and others
2.0	20/03/2024	Updates and appendices added
2.4	22/03/2024	Updated Maps and new section
2.6	25/03/2024	Final technical edits
2.7	11/04/2024	Final formatting edits

Approval history:

This document has been reviewed and approved by the following:

Name	Role	Date	Version
Roger Proudfoot	Project Senior Responsible Officer	03/04/2024	2.6

*This project was funded by the Department for Environment, Food and Rural Affairs (Defra) as part of the **marine Natural Capital and Ecosystem Assessment (NCEA) programme**. The marine NCEA programme is leading the way in supporting Government ambition to integrate natural capital approaches into decision making for the marine environment. Find out more at https://www.gov.uk/government/publications/natural-capital-and-ecosystem-assessment-programme*

Table of Contents

Ways to read this report

1. Key messages and highlights

There is a summary of Key Messages of the mNCEA Pelagic project monitoring programme showing where and when samples were collected at the beginning of this report. However, if you need to understand the development of the program then the **Overview of Sampling** in the **Year 1 Report** would also be useful (Marine Natural Capital and Ecosystem Assessment (mNCEA) Programme – Pelagic Monitoring Programme Project Year 1 Monitoring, programme, sampling analysis, and data collection).

2. Understanding methods

Full details of methods can be found in the **Year 1 report** and its appendices together with the sampling rationale for the Environment Agency, Cefas and Marine Biological Association (see Marine Natural Capital and Ecosystem Assessment (mNCEA) Programme – Pelagic Monitoring Programme Project Year 1 Monitoring, programme, sampling analysis, and data collection).

New data and methods from wider partners, if very different, are summarised in section 2, Overview of sampling.

3. Reviewing results

This report mainly focus on data *collection* in this year and a detailed statistical analysis would be inappropriate on less than 2 years' worth of data. Most of this data feeds into other work packages.

However, some results and analysis are presented on this section to illustrate the data and to reassure us that community patterns are as expected (for example, greatest numbers in the summer months and then tailing off).

The Plankton Lifeform Indicator Tool (PLIT) is also presented as a means of exploring the datasets. This is detailed in the Appendix.

4. Data storage

Each organisation has started at a different point in storing zooplankton data and the need to develop this. Each organisation has its own data management and curation processes. Ultimately all plankton data flows to the PLET (Plankton Lifeform Extraction Tool) hosted at the Data Archive for Marine Species and Seabed Habitats (DASSH) for storage, analysis, and public visibility.

PelCap - mNCEA NC34 Pelagic Programme: Year 2 Monitoring, sample analysis, data storage and data analysis report

Key Messages

Key messages on the monitoring and data aspects from Year 2 (2024/25) of the Pelagic Programme project ('PelCap'), delivered through the Marine Natural Capital Ecosystem Assessment (mNCEA) Programme:

- Continued targeted monitoring for zooplankton at nearshore, coastal, and offshore transects, enhancing data collection of zooplankton data
- New dataset discovery & utilisation – we now have available datasets from AFBI, PML, Newcastle, Isle of Man, SAMS, MSS
- Anecdotally, all participants report increasingly bad weather and port closures impacting sampling
- Historic datasets made useable
- Taxa diversity, abundance, and seasonality, generally in line with expectations
- First tranche of data exploration tools available
- Initial development of data management plans for program

1.Overview of Sampling

1. Year 2 objectives

1. Continuation with monthly nearshore zooplankton collection and identification from Environment Agency English inshore (Water Environment (Water Framework Directive) (England and Wales) Regulations, WER) water bodies – ranging between 20-25 water bodies with approximately 300 samples per full year, covering key inshore eco-hydrographic areas. The aim is to collect complete annual data sets with monthly resolution of zooplankton taxa that will be directly comparable to current co-located phytoplankton data sets, allowing comparison of inshore changes and impacts on food webs.

2. Continuation of the reinstated zooplankton monitoring at and around two Cefas English 'SmartBuoy' sites together with adjacent transects of supporting sites collecting water quality data. The aim is to collect annual data with (at least) monthly resolution of zooplankton. These sampling sites provide critical zooplankton data from Thames Estuary plume and Liverpool Bay plume, aiding the identification of impacts on key natural capital assets (shellfish and fisheries) within the riverine influenced areas. Monitoring surveys are collaborative between EA and Cefas, undertaken on EA Coastal Survey Vessels (CSVs *Thames Guardian* & *Mersey Guardian*) with Cefas sampling staff. Part of the annual sampling is also taken on the Cefas Research Vessel *Endeavour the Thames* and occasionally the AFBI vessel *RV Corystes* for the Mersey sites. This is part of a wider objective to understand and resolve the challenges of working across institutes to deliver more collaborative monitoring.

3. Routine Continuous Plankton Recorder (CPR) collection and analysis from reinstated English routes to cover all key areas. Annual data sets resolved at the frequency of ships of opportunity (weekly/monthly). Completes spatial coverage of England (offshore) waters. The reinstated MBA routes run alongside the existing CPR survey (part funded by Defra), and the combined dataset to be used in all assessment work. The route details are the:

    a. suspended extended B-route (since 2017) - this is the area that goes towards the Northeast approaches of the UK from the Atlantic, it is a key "early warning" area as it tends to be where we first pick up new or warmer water species moving upwards from lower latitudes

    b. CPR KC route - the only CPR route to go through Dogger Bank region, important for fisheries

    c. HE route - a southern North Sea route

    All data from these restarted routes will be immediately comparable to all historical CPR data.

4. Acquire and access other data sets and monitoring programs particularly if they can provide a long time series. This has included integration of historical datasets from the North Sea, where comparable with current monitoring data. The aim is to eventually have all historical and current data within the DASHH (Data Archive for Marine Species

and Seabed Habitats) system to allow assessment of lifeforms relative to the zooplankton data.

2. Monitoring summary

The maps below indicate the sites from where data was acquired from this year's monitoring (Figure 1). We also have access to data from PML L4 (and E1) buoy sites from the Western Channel Observatory (https://doi.org/10.5194/essd-15-5701-2023, see Figure 10). Moreover, we are utilising data from 4 sites in Scotland.

It is important to remember that this 'extra' gap filling data is set on top of routine monitoring programs (for example the EA WER Phytoplankton program), many of which are subject to financial restrictions.

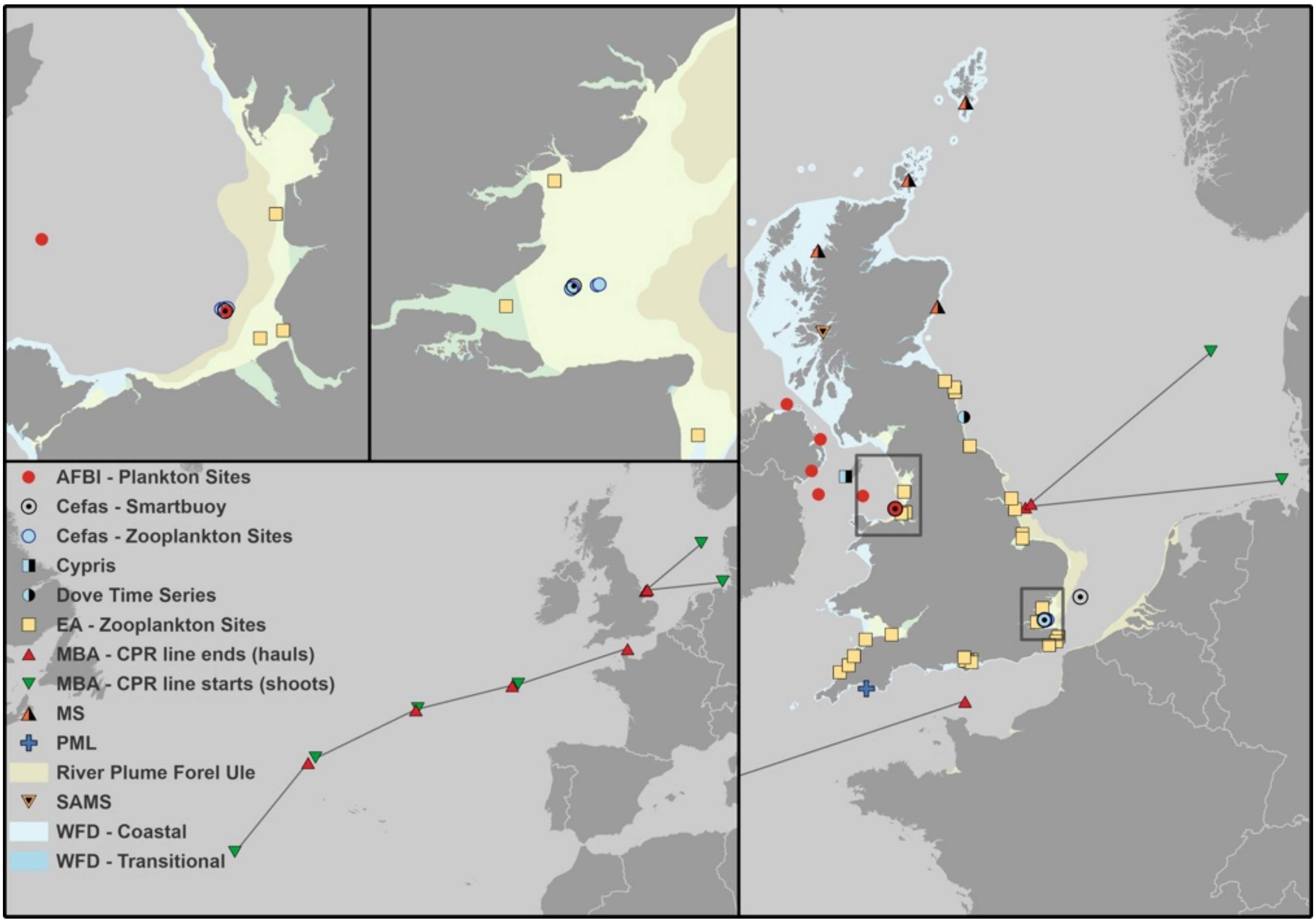


Figure 1. Year 2 mNCEA Pelagic Programme project, zooplankton monitoring sites and transects

2.1. Environment Agency

The Environment Agency's detailed site selection methodology can be found in in the Year 1 monitoring report. The final sites and some of their geographic reporting contexts are shown in Table 1 and Figure 3. Numbers in brackets are number of sites in that geographic area.

Regional Sea	EA Region	Water body	Site Name
Northern North Sea		Northumberland North (1)	OFFSHORE CHESWICK SANDS
		Farne Islands to Newton	FARNE ISLANDS 2.5KM OFF BEADNELL BAY (NORTHUMBERLAND WFD SITE 09)

North Sea (4)	North East (5)	Haven (2)	FARNE ISLANDS 2KM E OF INNER FARNE (THE BUSH) (FARNE ISLANDS WFD 03)
		Tees (1)	TEES AT REDCAR JETTY (SURFACE)
Southern North Sea (6)		Yorkshire South (1)	YORKS COAST - WITHERNSEA
	Anglian (4)	Lincolnshire (2)	LINCS COAST HAILE SAND FLAT S.YORKSHIRE LINCS.
			LINCS COAST OUTER DOGS HEAD 4.5 KM O/S
		Wash (1)	WASH SITE 33 - THE WELL/LYNN DEEPS 2 CONNECTED TO SPT WA560348 NEW SPT CREATED AS SITE MOVED
		Blackwater Outer (1)	VIRLEY CHANNEL OUTER R.BLACKWATER ST.PETER FLATS
	Thames (1)	Thames Lower (1)	THAMES AT NO.2 SEA REACH (77.6KMS BELOW LONDON BRIDGE)
Eastern Channel (8)	Southern (8)	Kent South (3)	GOODWIN FORK BUOY - INVESTIGATIONS BASELINE SURVEY
			SOUTH FORELAND - INVESTIGATIONS BASELINE SURVEY
			I KM SOUTH OF FOLKESTONE PIER, SOUTH KENT
		Portsmouth Harbour (1)	PORTSMOUTH HARBOUR MOUTH SAMPLING POINT
		Solent (3)	EAST BRAMBLES SAMPLING POINT
			RYDE-SHELLFISH WATER, 50'44.750N, 01'06.340W AT NE MINING GROUND BUOY
			COWES-SHELLFISH WATER, 50'46.380N, 01'17.500W AT PRINCE CONSORT BUOY
		Southampton Water (1)	FAWLEY SOUTH SAMPLING POINT
Western Channel & Celtic Sea (5)	South West (5)	Cornwall North (3)	NORTH CORNWALL OFF HARLYN BAY (WFD02)
			NORTH CORNWALL OFF BOSSINEY (WFD05)
			NORTH CORNWALL OFF SANDY MOUTH (WFD04)
		Barnstaple Bay (1)	BARNSTAPLE BAY OFF WOOLACOMBE (WFD 01)
		Bristol Channel Inner South (1)	INNER BRISTOL CHANNEL OFF MINEHEAD B (WFD 10)
Irish Sea (3)	North West (3)	Mersey Mouth (3)	MERSEY ESTUARY AT BUOY C21 HELICOPTER POINT 5
			COASTAL SURVEY NRA-173 BLACKPOOL: SITE SLC 40
			COASTAL SURVEY NRA-170 N WIRRAL: SITE SLC 55 & WLA 1

Table 1. Geographic regions and sites sampled under the Environment Agency sampling programme 2022

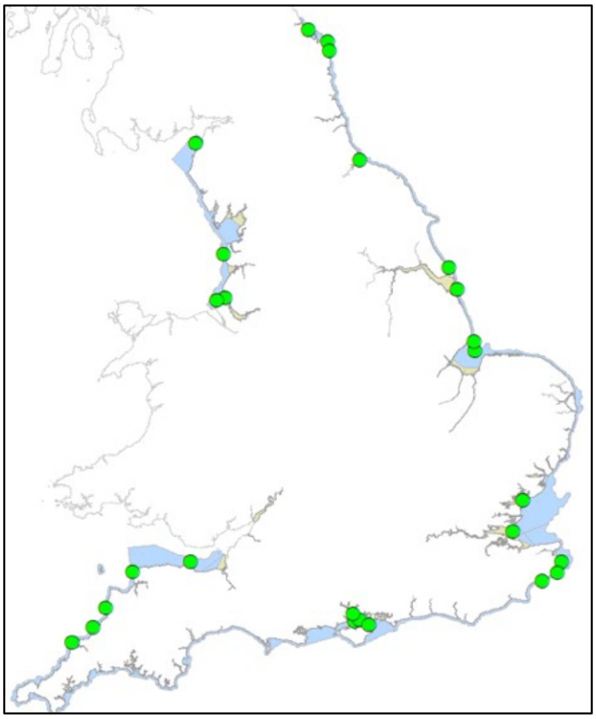


Figure 2. Environment Agency sampling sites

It was found that some sites had to be adjusted slightly for operational reasons. In addition, we are collecting an extra sample from the Solway for Natural England from September.

We can analyse the success of the sampling at various geographical scales. Over all regions, we can see that, as would be expected, the summer months were most successful (the dark mustard colour on the grand totals row). However, at the regional scale there are challenges in getting samples from both the Thames and the Southwest region which is mainly influenced by poor weather conditions and mechanical issues.

Row Labels	2022 Jun	Jul	Aug	Sep	Oct	Nov	Dec	2022 Total	% Months samp	2023 Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	2023 Total	% Months samp	Grand Total
NEast			8	4	4	7	1	24	100%	4	1	5	3	4	4	3	2	4	4	5	39	100%	63
Anglian			10	4	2	5	4	25	100%	2	5	5	2	5	5	5	5	5	3	2	44	100%	69
Thames			2		1	1		4	60%						1	1	1	1	1	1	6	55%	10
Southern	2	6	10	6	4	2	4	34	100%	3	4	8		7	7	7	7	7	8	7	65	91%	99
SWest			6	3	6		1	16	80%		4	1	5	5	5	1		5			26	64%	42
NWest			3	3	3	3	3	15	100%		3	3	3	3	3	3	3	4	3	4	32	91%	47
Grand Total	2	6	39	20	20	18	13	118		9	17	22	13	24	25	20	18	26	19	19	212		330

Table 2. Regional level analysed Environment Agency zooplankton samples. Regions organised clockwise around England from the North-East to North-West. Numbers represent fully analysed samples per month. Orange area represents the 2022 testing period. For totals per month and % samples, darker shading represents high numbers.

When we resolve this data further to water body level, both the Southeast corner (from Essex to Kent) and the Northern coast of the Southwest Peninsula pose problems for sampling (Table 3).

Region	Water body	2022 Jun	Jul	Aug	Sep	Oct	Nov	Dec	2022 Total	% Months samp	2023 Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	2023 Total	% Months samp	Grand Total
NEast	Northumberland North			1	1	1	2		5	80%	1		1	1	1	1		1	1	1	1	9	82%	14
	Farne Islands to Newton Haven			4	2	2	4		12	80%	2		2	2	2	2	2	1	2	2	2	19	91%	31
	TEES			3	1	1	1	1	7	100%	1	1	2		1	1	1		1	1	2	11	82%	18
Anglian	Yorkshire South			1	1	1	1	1	5	100%	1	1	2		1	1	1	1	1	1		10	82%	15
	Lincolnshire			5	2		2	2	11	80%		2	2	1	2	2	2	2	2	1		16	82%	27
	Wash Outer			2	1		1	1	5	80%	1	1	1	1	1	1	1	1	1			9	82%	14
	Blackwater Outer			2		1	1		4	60%		1			1	1	1	1	1	1	2	9	73%	13
Thames	THAMES LOWER			2		1	1		4	60%						1	1	1	1	1	1	6	55%	10
Southern	Kent South			2	2		2		6	60%		3			3	3	3	3	3	3	3	24	73%	30
	Solent	2	6	6	3	3		3	23	80%	3		6		3	3	3	3	3	3	3	30	82%	53
	SOUTHAMPTON WATER			2	1	1		1	5	80%		1	2		1	1	1	1	1	2	1	11	82%	16
SWest	Barnstaple Bay			2	1	1			4	60%		1		1	1	1			1			5	45%	9
	Cornwall North			4	1	3			8	60%		3		3	3	3	1		3			16	55%	24
	Bristol Channel Inner South				1	2		1	4	60%			1	1	1	1			1			5	45%	9
NWest	Mersey Mouth			3	3	3	3	3	15	100%		3	3	3	3	3	3	3	3	3	3	30	91%	45
	Solway Outer South																		1		1	2	18%	2
	Grand Total	2	6	39	20	20	18	13	118		9	17	22	13	24	25	20	18	26	19	19	212		330

Table 3. Water body level analysed Environment Agency zooplankton samples. Water body names in CAPITALS indicate estuarine water bodies. Numbers represent fully analysed samples per month. Orange area represents testing period. Percentage of months successfully sampled based on 5 months for 2022 and 11 months for 2023. For totals per month and percentage (%) samples, darker shading represents high numbers.

Resolving further to the site level, we can see that some sites are (or have been) particularly challenging to sample, especially in the Southwest and we may need to consider relocating these sites to areas where there is less susceptibility to weather to have reliable long-term monitoring (Table 4).

Table 4. Site level analysed Environment Agency zooplankton samples. Water body names in CAPITALS indicate estuarine water bodies. Numbers represent fully analysed samples per month. Area represents testing period. Percentage of months successfully sampled based on 5 months for 2022 and 11 months for 2023. For totals per month and percentage (%) samples, darker shading represents higher numbers.

Region	Water body	Site codes	2022							2022	% Months	2023											2023	% Months	Grand
			Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total	samp	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Total	samp	Total
NEast	Northumberland North	NNC005N			1	1	1	2		5	80%	1		1	1	1	1		1	1	1	1	9	82%	14
	Farne Islands to Newton Haven	FNH002N			2	1	1	2		6	80%	1		1	1	1	1	1		1	1	1	9	82%	15
		FNH004N			2	1	1	2		6	80%	1		1	1	1	1	1	1	1	1	1	10	91%	16
	TEES	TEE002N			3	1	1	1	1	7	100%	1	1	2		1	1	1		1	1	2	11	82%	18
Anglian	Yorkshire South	SYK006N			1	1	1	1	1	5	100%	1	1	2		1	1	1	1	1	1		10	82%	15
	Wash Outer	OWS003N			2	1		1	1	5	80%	1	1	1	1	1	1	1	1	1			9	82%	14
	Lincolnshire	SYK003N			3	1		1	1	6	80%		1	1		1	1	1	1	1	1		8	73%	14
		SYK005N			2	1		1	1	5	80%		1	1	1	1	1	1	1	1			8	73%	13
	Blackwater Outer	OBW005N			2		1	1		4	60%		1			1	1	1	1	1	1	2	9	73%	13
Thames	THAMES LOWER	TMT005N			2		1	1		4	60%						1	1	1	1	1	1	6	55%	10
Southern	Kent South	SKT002N				1		1		2	40%		1			1	1	1	1	1	1	1	8	73%	10
		SKT003N											1			1	1	1	1	1	1	1	8	73%	8
		SKT004N			2	1		1		4	60%		1			1	1	1	1	1	1	1	8	73%	12
	Solent	SOL001N		2	2	1	1		1	7	80%	1		2		1	1	1	1	1	1	1	10	82%	17
		SOL002N	2	2	2	1	1		1	9	80%	1		2		1	1	1	1	1	1	1	10	82%	19
		SOL005N		2	2	1	1		1	7	80%	1		2		1	1	1	1	1	1	1	10	82%	17
	SOUTHAMPTON WATER	SOU001N			2	1	1		1	5	80%		1	2		1	1	1	1	1	2	1	11	82%	16
SWest	Cornwall North	NCW002N			2	1	1			4	60%		1		1	1	1			1			5	45%	9
		NCW004N			2		1			3	40%		1		1	1	1			1			5	45%	8
		NCW005N					1			1	20%		1		1	1	1			1			5	45%	6
		NNC005N																1					1	9%	1
	Barnstaple Bay	BRB001N			2	1	1			4	60%		1		1	1	1			1			5	45%	9
	Bristol Channel Inner South	IBS002N				1	2		1	4	60%			1	1	1	1			1			5	45%	9
NWest	Mersey Mouth	LIV002N			1	1	1	1	1	5	100%		1	1	1	1	1	1	1	1	1	1	10	91%	15
		LIV003N			1	1	1	1	1	5	100%		1	1	1	1	1	1	1	1	1	1	10	91%	15
		LIV006N			1	1	1	1	1	5	100%		1	1	1	1	1	1	1	1	1	1	10	91%	15
	Solway Outer South	OSS006N																		1		1	2	18%	2
		Grand Total	2	6	39	20	20	18	13	118		9	17	22	13	24	25	20	18	26	19	19	212		330

Ongoing and increasing poor weather seems to be a common feature for disruptions of sampling. This has been reported by other participants where the increasing occurrence of poor weather has caused delays, postponed sampling, or limited harbour access.

2.2. Cefas

Cefas continued to sample at and around the SmartBuoy sites as per year 1 activity (see Figure 3, and year 1 monitoring report) with additional water quality (WQ) monitoring conducted by Cefas RV *Endeavour* and in the transects from inshore to offshore on the EA survey vessels. Additional data sets have also been collected from the Grant In Aid monitoring programs and these are summarised in Figure 3 and Table 5.

In addition, historical and current sampling data were compiled into accessible and published datasets. A summary of data descriptions, data stewards and target databases are outlined in Table 5.

Data name	Description	Steward	Target
mNCEA Zooplankton 200um ring net samples	Samples collected at Mersey and Thames estuary by ring net and analysed by microscopy	James Scott	PLET, DASSH and Datahub (using the Darwin format)
mNCEA Water Quality data	Samples collected at Mersey and Thames estuary through various methods.	Paul Nelson & Elise Brabben	BODC
mNCEA Phyto bottle samples (limited)	Samples collected at Mersey and Thames estuary by ring net and analysed by microscopy.	Amy Mace	PLET and Datahub
Historic samples (5 case studies)	Data from the Cefas 'archives' to try to populate areas of low data on PLET (e.g. mid north sea)	Suzanne Painting	PLET and Datahub
SmartBuoys phytoplankton update	Update to existing data in PLET tool and Cefas datahub. Phyto samples analysed by microscopy.	Amy Mace	4 year UPDATE PLET and Datahub

Gabbard data	Three-year update of monthly samples	Sophie Pitois	3 year UPDATE PLET and Datahub

Table 5. Cefas data sets related to NC34 Pelagic Programme

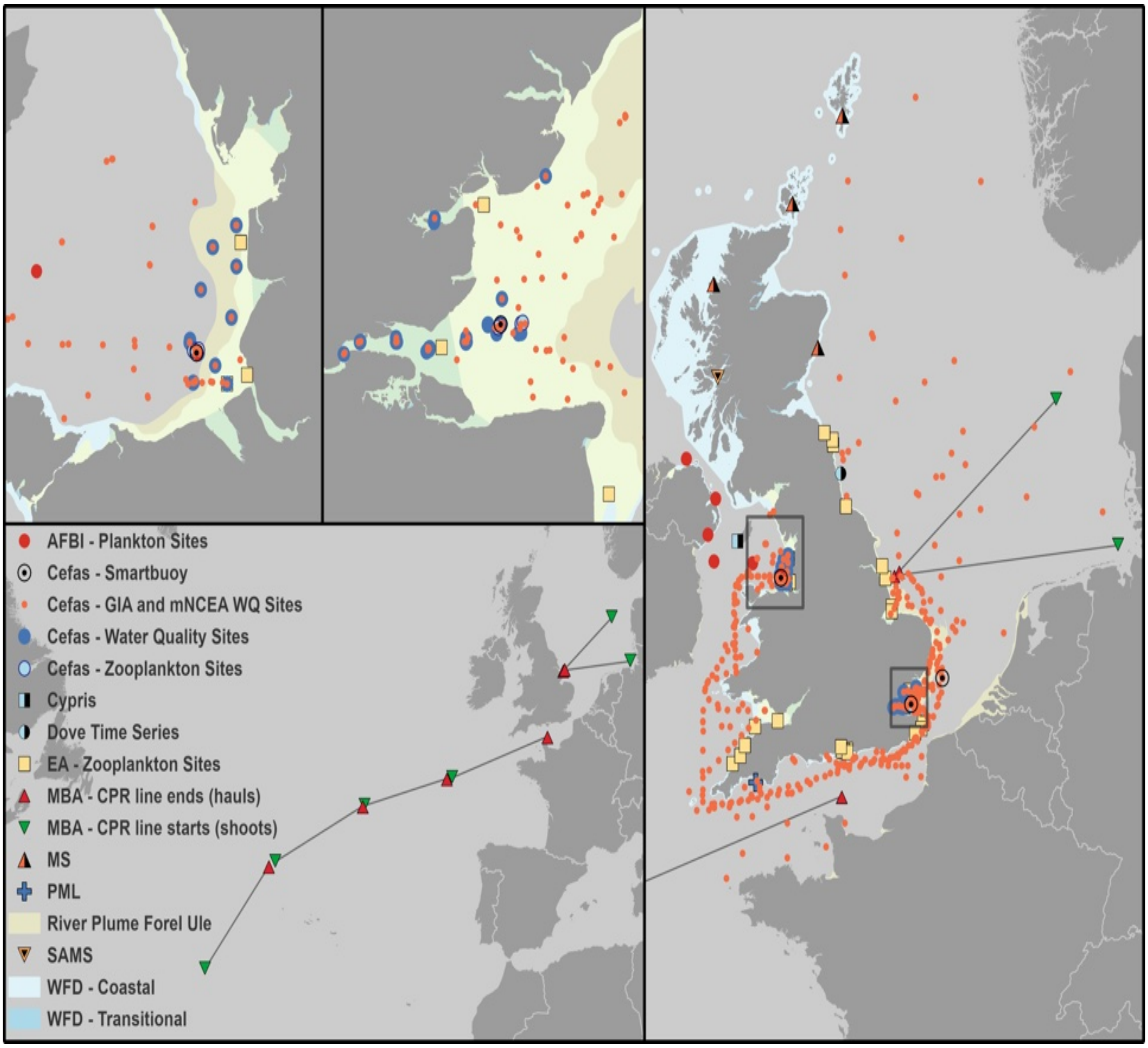


Figure 3. Monitoring sites including Cefas water quality (WQ) sites. The open circles with a dot indicate SmartBuoy locations. Blue closed circles: zooplankton & supporting data sites; Orange dots: water quality sites.

The success of sampling is similar to the inshore sampling program (EA) when the smaller coastal survey vessels were used with 3 months of sampling being impacted by bad weather and vessel logistics. However, that equates to only losing one month in Thames and two months in the Mersey. Figure 4 illustrates both method development on the smaller EA vessels, and the number of times the sampling was supported by either the RV *Endeavour* (Cefas) or RV *Corystes* (AFBI).

Method development on the smaller boats can be seen with the introduction of flow cytometry samples in the early months of the year and for new technologies including Coloured Dissolved Organic Matter (CDOM) and Forel Ule (Ocean Colour).

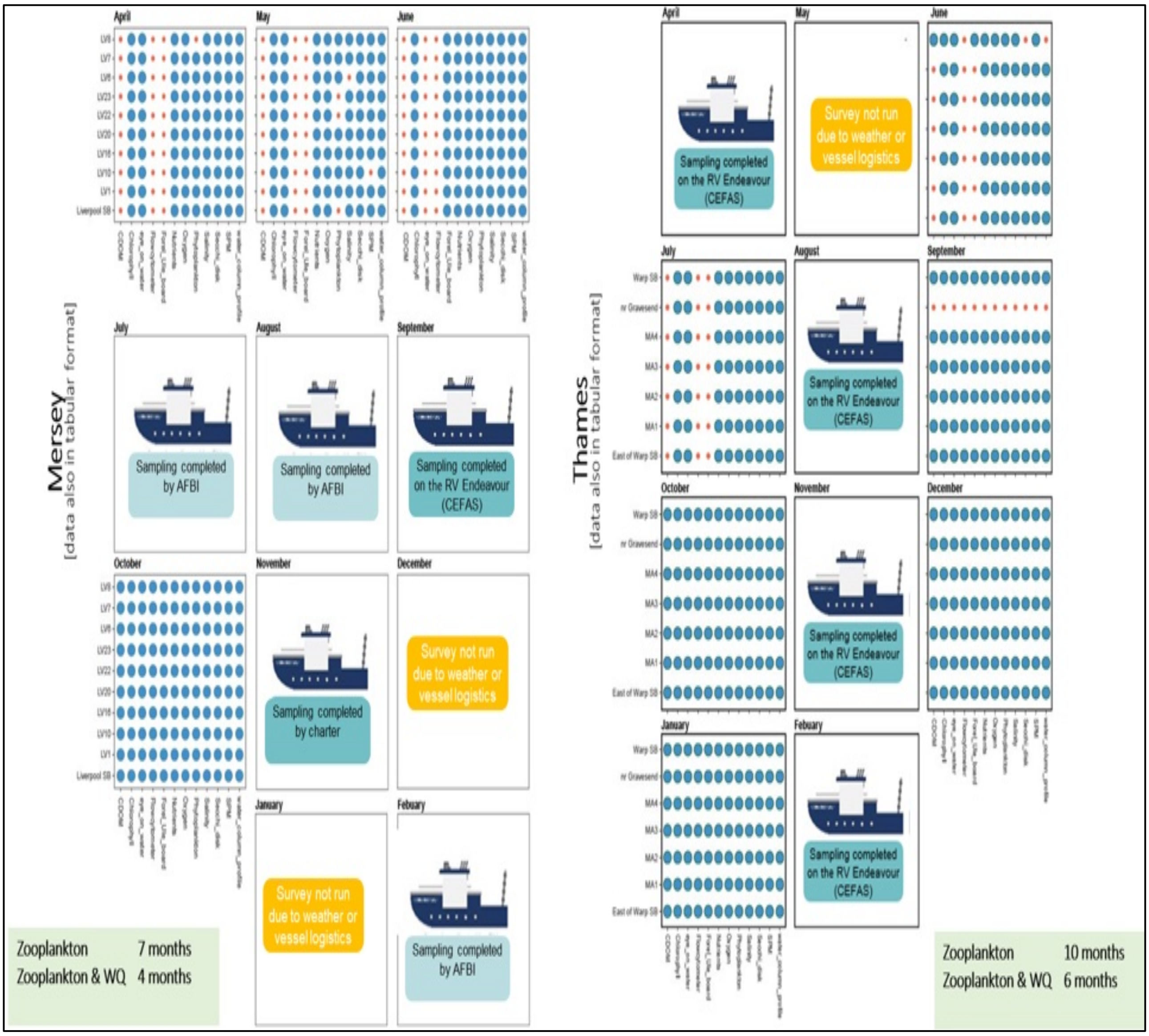


Figure 4. Success of increased sampling in the Thames and Mersey supported by EA survey boats, Cefas RV Endeavour and AFBI research vessels. The blue and orange dots represent the water quality and plankton samples that are taken on the EA survey boats, with orange dots in the earlier months highlighting what samples were not collected, and blue dots where the sample had been successfully taken. The move from orange to blue dots in later months showing the success of adding new parameters to the EA surveys including CDOM, flow cytometry and ocean colour.

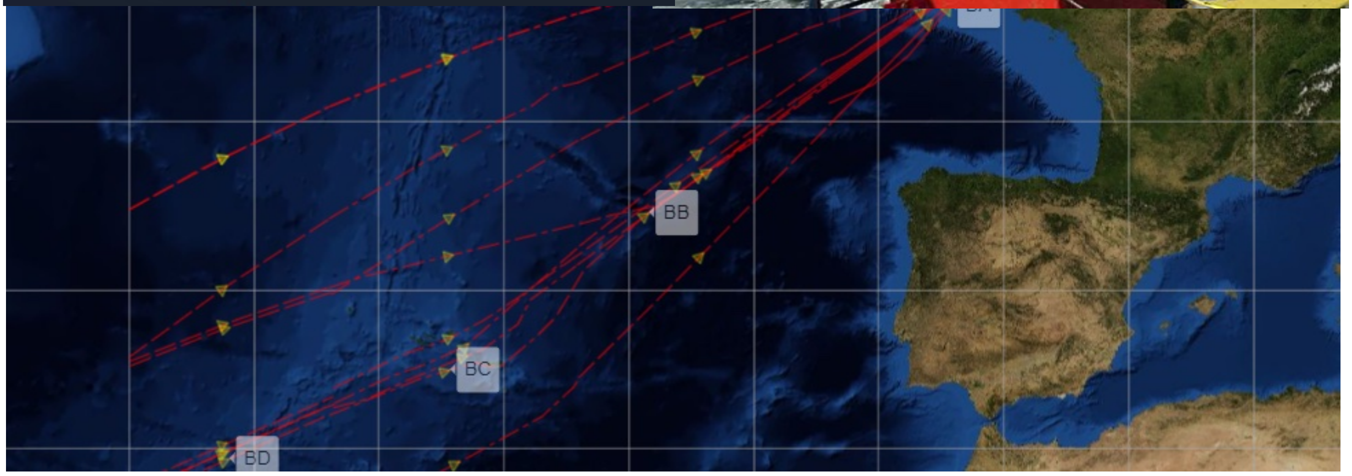
A larger version of Figure 4 can be found in Appendix 1.

2.3. MBA - mNCEA funded CPR routes

May 2022 saw several routine CPR English routes reinstated under the mNCEA programme. These have been continued. These routes are from the NE Atlantic to Southwest Approaches (BB, BC, BD), Dogger Bank (KC) and Southern North Sea (HE). Some of the routes are constrained in enclosed waters and the tracks are almost identical, but routes in the open ocean are much more variable.

Figure 5. Continuous Plankton Recorder (CPR) tows: 569BD, 569BC, 595BB, 609BA, 52KC, 372HE towed regularly over 2022/23

Figure 6. Left: CPR, Right: CPR about to be deployed in the water



The KC route is operated by the vessel *SC Connector*, and part of Norway's fleet, running from Immingham to the Norwegian coastline.

The HE route, undertaken by the vessel *Selandia Seaways* (Figure 7) runs from Immingham to the German coastline. Both of these transects operate across some important fisheries areas.

The BB, BC, BD tows is operated currently by the cargo ship *Lombok Strait*, which operates between the UK (English Channel region) and Barbados. Transecting across shallower UK waters to deeper mid Atlantic waters, this is an important area, as changes in the plankton community can give early warning regarding potential changes in UK waters.

Figure 7. Sealandia Seaways © Florian Stelling.

3. Monitoring summary - new data sets

A number of new datasets have been discovered and added to the plankton Lifeform Extractor Tool (PLET) database (Table 6). Full tables can be found in Appendix 1.

Institute	Dataset
Agri-Food and Biosciences Institute (AFBI)	Western Irish Sea Long term monitoring
Centre for Environment, Fisheries and Aquaculture Science (Cefas)	Cefas SmartBuoy Marine Observational Network - UK Waters Phytoplankton Data 2001-2019
	Cefas West Gabbard zooplankton abundance time series monitoring using ZooScan from 2016 to 2020
Environment Agency (EA)	EA CHL 2000-2020
	EA PHYTO 2000-2020
Isle of Man Government (IOM)	Cypris Station Phytoplankton Abundance
Marine Biological Association (MBA)	Continuous Plankton Recorder
Marine Scotland (MS)	MSS Loch Ewe Phytoplankton
	MSS Loch Ewe zooplankton
	MSS Scalloway Phytoplankton dataset
	MSS Scapa Phytoplankton dataset
	MSS Stonehaven Chlorophyll data
	MSS Stonehaven Phytoplankton
	MSS Stonehaven zooplankton
Natural Resources Wales (NRW)	NRW WFD Phytoplankton classification data 2007_2019
Newcastle University (NU)	Dove Time Series Ichthyoplankton zoo
	Dove Time Series WP2 Microsope and FlowCam
	Newcastle University Dove Time Series WP2 and Horizontal WP3
	Newcastle University Dove Time Series WP2 FlowCam zoo
	Newcastle University/Cefas Dove Time Series P200
Plymouth Marine Laboratory (PML)	PML_L4 chl a
	PML_L4 phytoplankton
	PML_L4 zooplankton
Scottish Association for Marine Science (SAMS)	SAMS-LPO-Phyto-Dec2021

Table 6. Summary of the current and new datasets collected that have been added to PLET.

3.1. Agri-Food and Biosciences Institute (AFBI) Plankton data set

As of early 2023, AFBI takes samples of phytoplankton and zooplankton from a total of five stations, approximately 10 times a year, and samples at Cefas Liverpool Bay mooring approximately 4 times a year (see Figure 8).

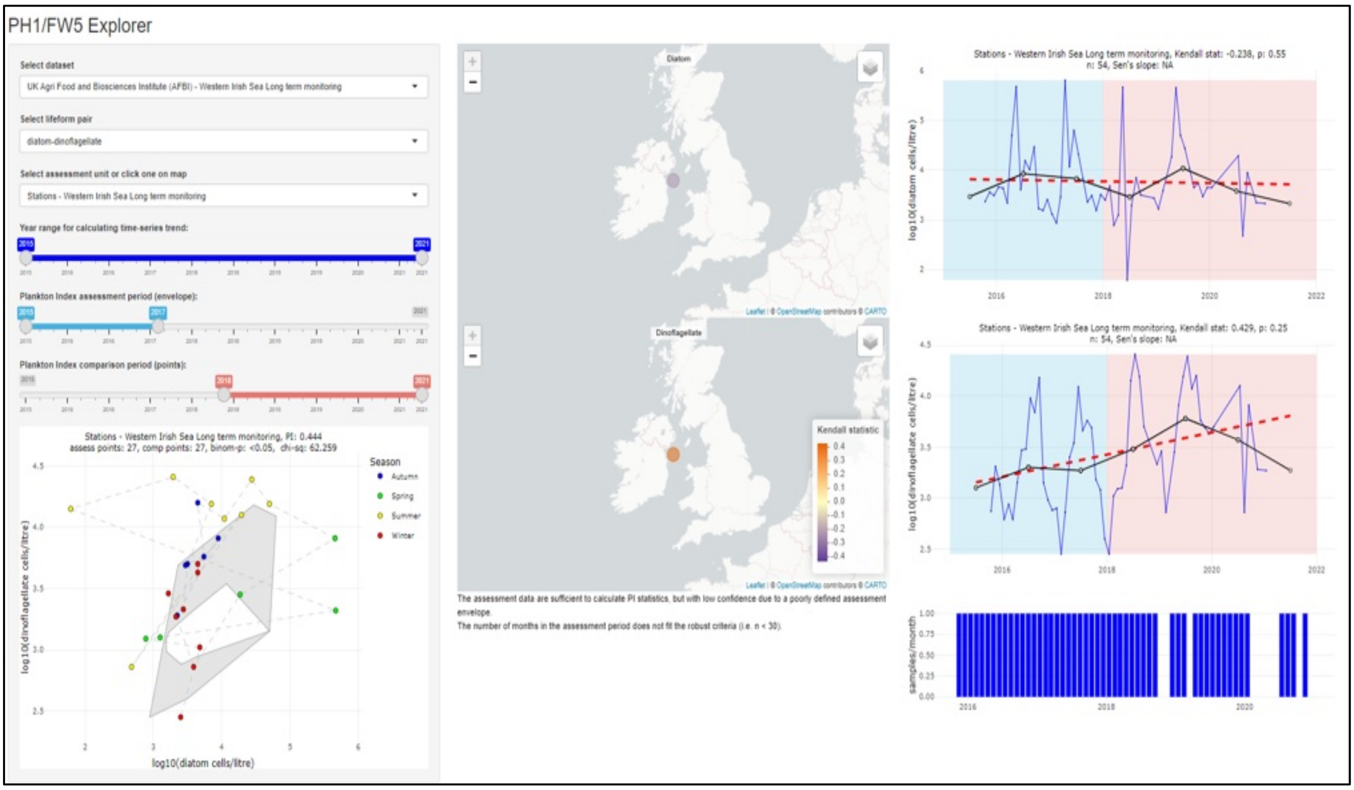
Together with some historical data our data set consists of:

- 5 years (2016-2020) historic phytoplankton data submitted to PLET, includes AFBI's historic holdings from **one** of the Irish Sea stations (see data summary infographic, Figure 9)
- 3 years (2015-2017), 26 zooplankton samples from the AFBI Western Irish Sea station (38a)
- 3 years (2015-2017), 22 zooplankton samples from AFBI LB06 station
- 3 years (2015-2017), 23 zooplankton samples from the Cefas SmartBuoy site in Liverpool Bay.

Processing is ongoing for zooplankton samples from 2018, and phytoplankton from 2021 with ongoing submission to PLET.

Figure 8. AFBI Irish sea monitoring stations

Figure 9. AFBI data block summary for site WF85, illustrating trends in diatoms and dinoflagellates together with sampling frequency.



3.2. PML Southwestern Channel Observatory

The Western Channel Observatory (WCO) comprises a series of pelagic, benthic, and atmospheric sampling sites within 40 km of Plymouth, UK, that have been sampled by the Plymouth institutes on a regular basis since 1903. The key data sets for this project are mainly from L4.

The features of all the WCO data are detailed in a recent publication:

[ESSD - The Western Channel Observatory: a century of physical, chemical, and biological data compiled from pelagic and benthic habitats in the western English Channel (copernicus.org)](#).

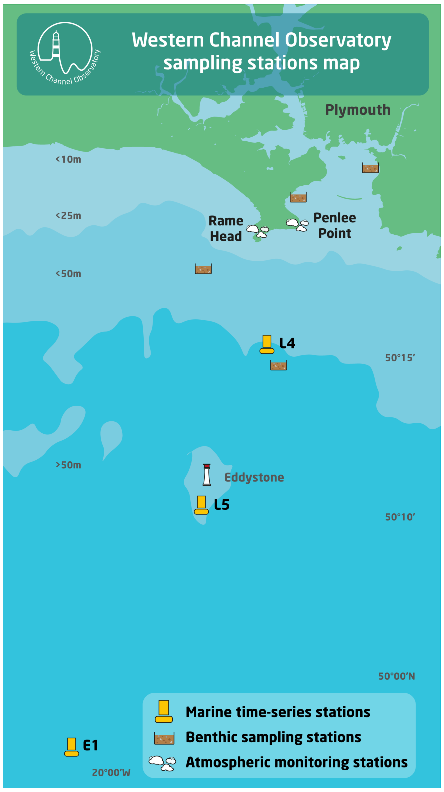


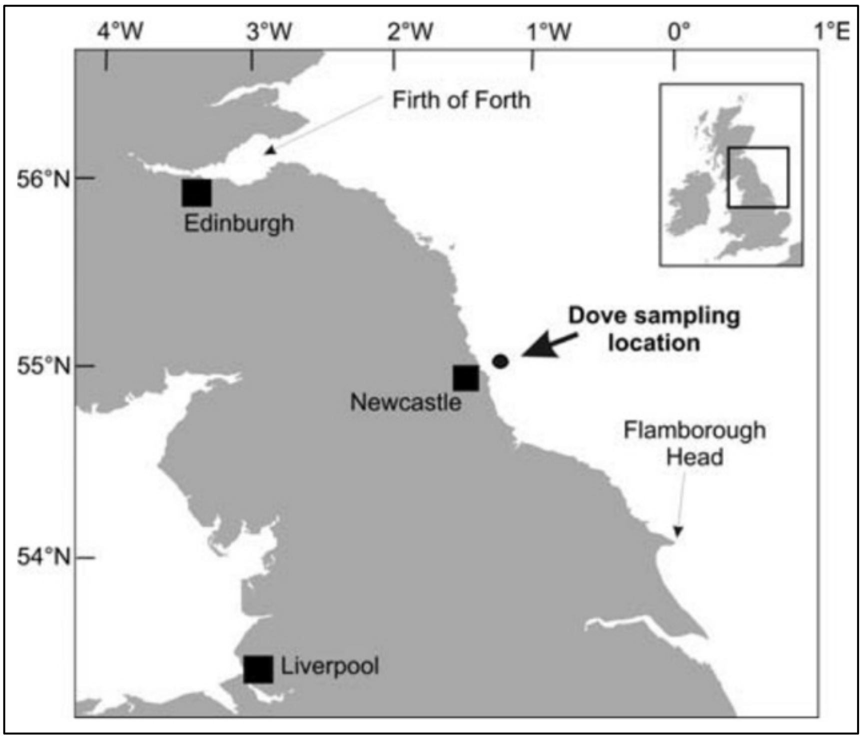
Figure 10. Western Channel observatory sites. Accessed from **Western Channel Observatory**. Key data for this project are mainly from L4.

3.3. Newcastle university Dove Maine Laboratory time series

This time series was established in **1968.** Monthly sampling occurs at a site of ~ 54m depth, 5.5 nautical miles east of Blyth port (Northumberland) (55^{0} 07'N, 01^{0} 20'W).

Figure 11. Dove time series location.

Three nets (UNESCO standard WP2, WP3 & P200) are used in conjunction with a flowmeter in four ways:



- **WP3** (1mm mesh) towed horizontally for 10 minutes at ~ 25m depth.
- **WP3** (1mm mesh) vertical haul (50m to surface)
- **WP2** (200µm mesh) hauled vertically (50m to surface), 4 times with each haul combined.
- **P200** (63µm mesh) hauled vertically (50m to surface), 4 times with each haul combined.

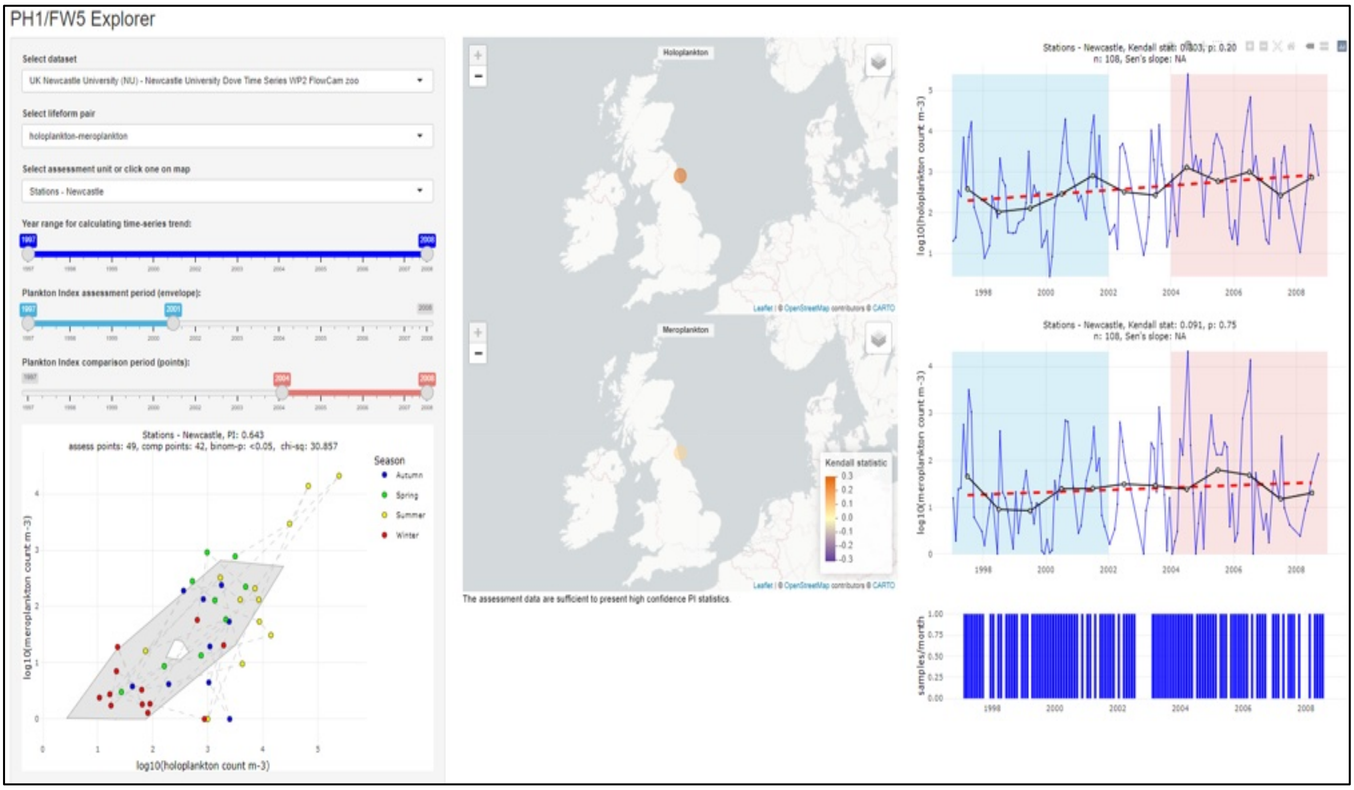
The WP2 and P200 nets were hauled from approximately 50 m, four times and collated to cover a net distance of 200 m and a net volume of 50 m3. The samples are fixed in 4% buffered formaldehyde, and subsequently transferred to 70% IDA for long – term storage. The physical samples are stored at the Dove Marine Laboratory

The time series was established as a zooplankton monitoring programme, and the methodology reflects this. However, the preserved P200 samples were analysed semi-quantitatively (M Baptie, 2013) to determine long term changes in phytoplankton community composition. Recently, flow cam methods (image analysis) have also been investigated. An example of the data is show in the data block in Figure 12.

Figure 12. Data block for Dove Zooplankton WP2 Flow Cam. Illustrating trends in holoplankton and meroplankton together with sampling frequency

3.4. Isle of Man Cypris Station

The Cypris station located approximately 2.5nm (approx. 5km) due west of Port Erin Bay was adopted as an offshore monitoring station in 1954 (Figure 13). The Cypris data have been collected at frequencies ranging from weekly to monthly depending on season, boat availability and weather. Following the closure of the Port Erin Marine Laboratory in 2006, the Manx Government Laboratory took responsibility for the routine collection and analysis of marine data from 2006 to 2017. As of 2021, the responsibility for marine monitoring moved to the Environmental Protection Unit. The Isle of Man restarted their marine



monitoring program at the Cypris monitoring station in May 2023, when a 1.9m marine monitoring Fulmar buoy was deployed (Figure 13).

Since then, routine sampling trips have been undertaken roughly every three weeks to collect physical samples for phytoplankton, zooplankton, and nutrient analysis. The parameters the buoy monitors are: temperature, conductivity, salinity, turbidity, dissolved oxygen, pH, partial pressure of carbon dioxide (pCO2), chlorophyll, and wave height and direction.

Real time data from the buoy is automatically uploaded to a server that is publicly available online (app.konectgds.com/kiosk/d66ee3b7-36a2-4d7a-8545-b2f37313db70). Phytoplankton data is updated on the Isle of Man Marine monitoring website (https://www.gov.im/iommarinemonitoring#accordion) every three weeks or so. Zooplankton data will also be uploaded here when available.

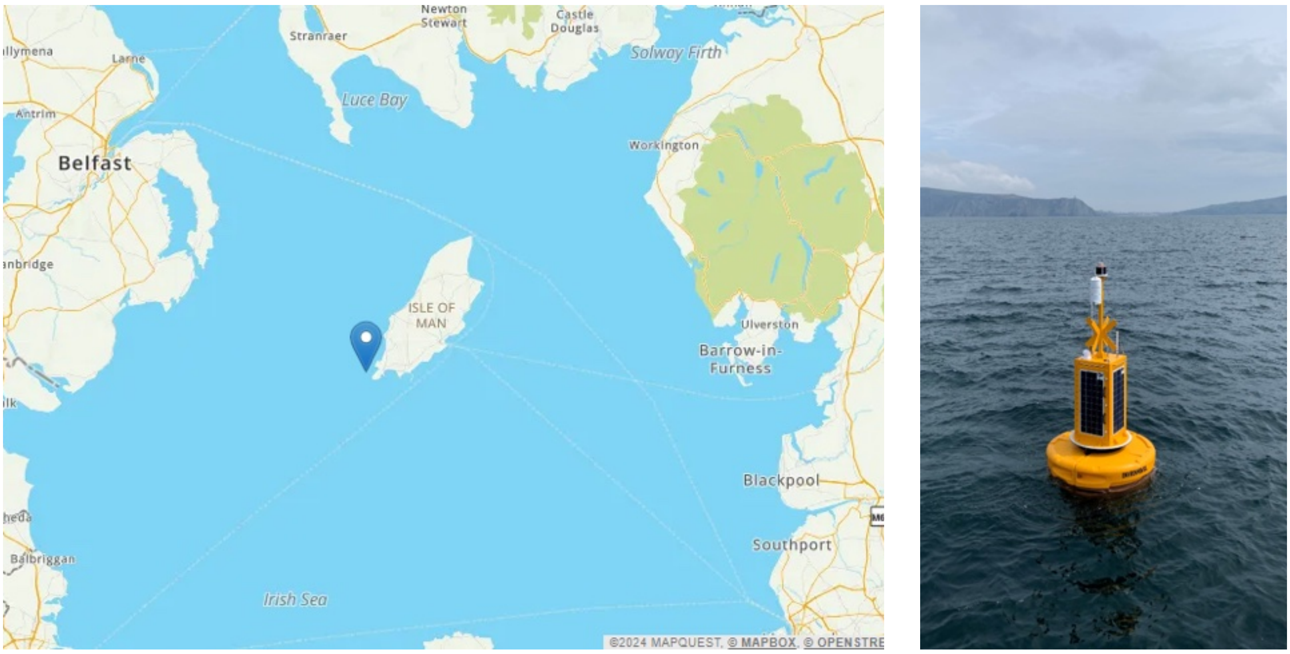


Figure 13. Above left: Location Cypris station, Above right: Cypris buoy.

3.5. Potential for Welsh zooplankton data

We consulted with Defra and Natural Resource Wales (NRW) to discuss the possibility of extra zooplankton samples being taken alongside some of their water quality sites; discussion are ongoing.

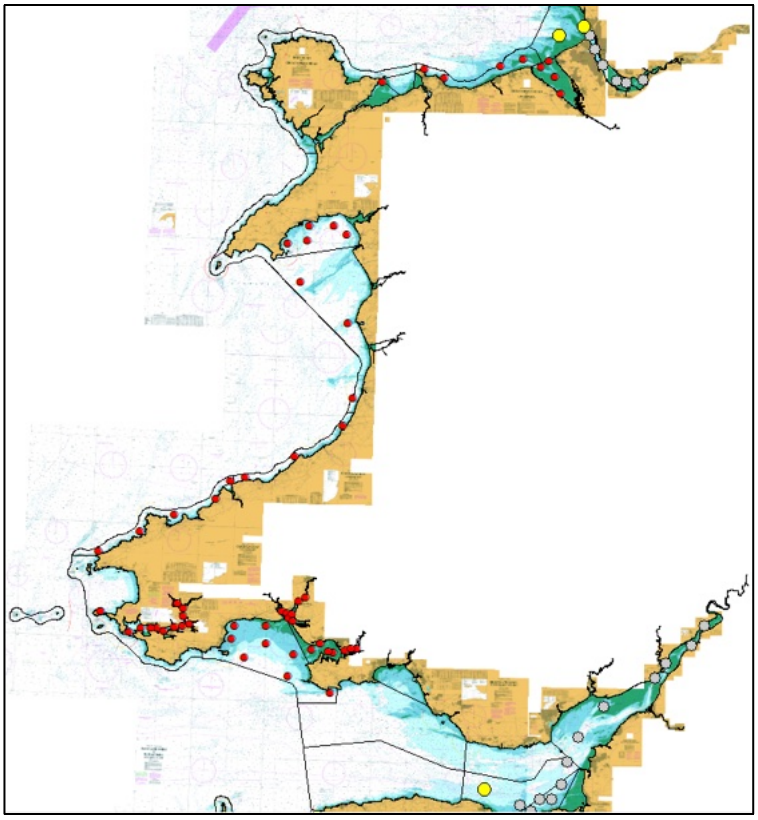


Figure 14. Water quality and phytoplankton sites sampled by the Environment Agency on behalf of Wales

4. Historical data

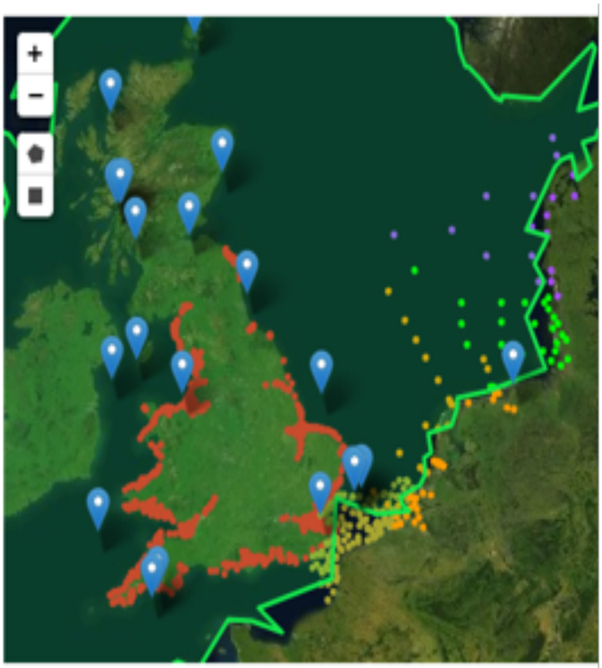
All organisations have historical data in various available or unavailable formats. Cefas have a wealth of historic plankton data in various digital archives. Some of these data date back over 80 years some are digitised from paper records. As part of the mNCEA project we have tried to deliver some of these data in a format appropriate for integration into existing plankton data repositories, mainly the Plankton Life Form Extraction tool (PLET) tool.

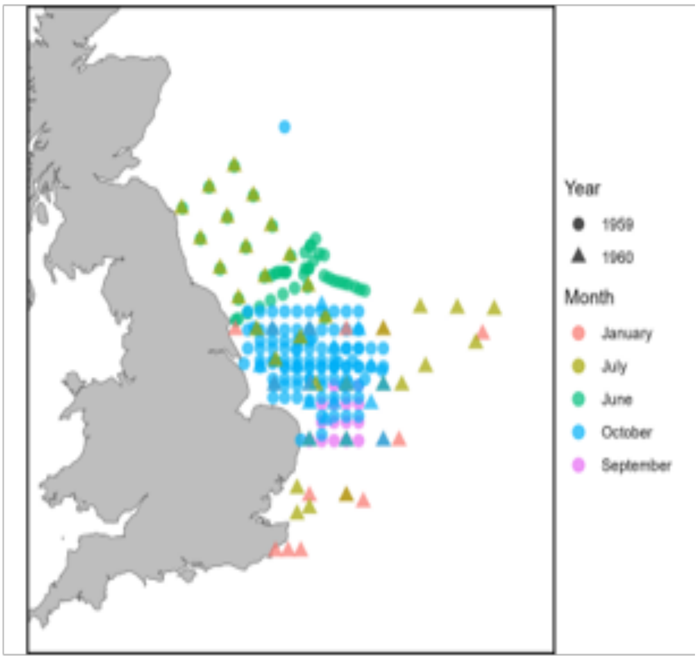
A few years ago, a report was produced which explored what data was available within Cefas (Figure 15).

Figure 15. Current data in the Plankton Life Form Extraction tool (PLET) show in the left panel. On the right is some example data from the Cefas archive. The data shown on the right panel will be delivered as part of the mNCEA PELCAP project.

From this report, 7 example data sets were chosen (Table 7, Figure 15).

These were used to test a code-based method for restoring the data and presenting it in a format compatible with the PLET tool. These datasets are show next to the current data in the PLET tool in Figure 15. It is apparently there are some large gaps in the North Sea that the Cefas historic data may fill.





Restoring the data is not straightforward. Often parts of the metadata required are not with the associated data and finding the corrected information can be a time-consuming process. The methods associated with the data, required to use the data ecologically can also be absent. These few case studies demonstrate that it is possible to restore the data but at considerable resource cost.

Dataset name	Sea area	Net	Details of Sample Analysis	Year	Stations	PLET Cleaned?
NorthSea_1959_SLAN_09	North Sea	Vertical trawl Hensen net tow <u>Assumed</u> 270 µm	Zooplankton identified, staged and counted. Light microscopy	1959	50	Yes
NorthSea_1959_SLAN_14					19	
NorthSea_1959_SLAN_14B					5	
NorthSea_1960_SLAN_01				1960	21	
NorthSea_1959_SLAN_14					**115**	
NorthSea_1959_SLAN_11					33	
ecosystem_connections		50 µm ring net	Zooplankton identified, sexed, staged and counted. Light microscopy	2007	3	

Table 7. Historic data that has been restored to be delivered as part of the mNCEA PELCAP project from Cefas data archives

5. Methods and method changes

Methods are detailed in the Year 1 Monitoring report ([Marine Natural Capital and Ecosystem Assessment (mNCEA) Programme – Pelagic Monitoring Programme Project Year 1 Monitoring, programme, sampling analysis, and data collection (wordpress.com)]) – Section 4 and Appendices. The basic methods have not changed for this year (apart from some minor site changes), but additional methods have been added to some surveys such as imaging (such as FlowCam), flow cytometry for picoplankton and molecular methods. Appendix 2 reviews some novel methods (further investigations into these will continue into year 3). See Table 7.

Institute	Phyto / Zoo	Collection method and analysis method	Sampling type	Location	Temporal coverage
Cefas	Zooplankton	Ring net vertical haul	Discreet	Celtic sea	11 years to present
		Zooscan and Ecotaxa		Every autumn	
		Plankton Imager	Continuous	Celtic sea	4 years to present
		Machine learning (in house)		Every autumn	
		Plankton Imager	Continuous	Celtic sea	3 years to present
		Machine learning (in house)		Every spring	
		Ring net	Discreet	Gabbard - outer Thames	8 years to present
		Zooscan and Ecotaxa		Target monthly	
Newcastle University	Zooplankton	Ring net	Discreet	Coastal North Sea	1968-present
		COI gene sequencing		Monthly	
SAMS	Phytoplankton	Imaging FlowCytobot	Continuous	Shetland	2.5 years Scalloway 1 year Cole Deep
		Machine learning (in house)			
PML	Zooplankton	Ring nets	Discreet	L4	From 2012
		FlowCam, Manual and Ecotaxa		Weekly samples	
		Water sample	Continuous	L4	From 2012 - mid-2023 exists
		18S V9 rRNA gene sequencing		Weekly samples	
	Zooplankton	Ring nets	Discreet	Ad hoc research cruises	

Exeter Uni	Zooplankton	FlowCam Macro		Ad hoc research cruises	
	Phytoplankton	Water sample	Discreet	Ad hoc research cruises	
		FlowCam 8400			
MBA	Zooplankton	Water sample	Discreet	Northeast Atlantic	From 2016 to 2019
		Machine learning (in house)		Sporadically	
MBA	Both	iCPR Holographic camera	Continuous	Northeast Atlantic	To start
		Custom algorithms	(every 5 sec)	Monthly	
PML	Zooplankton	Plankton Imager	Continuous	L4	To start
		Machine learning (in house)		Hourly	
	Phytoplankton	Imaging FlowCytobot	Continuous	L4	
		Machine learning (in house)		Minutely	
Scottish MD	Zooplankton	Ring nets	Discreet	Stonehaven / Loch ewe	To start
		FlowCam		Weekly samples	
EA	Phytoplankton	Imaging Flow Cytometer	Discreet	WER water bodies	To Start

Table 8. Current use of imaging and DNA within the PELCAP group. Where Cefas = Centre for Fisheries and Aquaculture science. SAMS = Scottish Association for Marine Science. PML = Plymouth Marine Laboratory. MBA = Marine Biological Association. MD = Marine Scotland. EA = Environment Agency.

2.Summarised results

The annually collected data sets are most useful when incorporated into ongoing time series or to extend the geographic range of other data sets. Hence this report has limited details in this area.

2.1. Plankton Lifeform Indicator Tool (PLIT)

A web-based tool (https://uk-pheg.shinyapps.io/ph1_shiny/) was developed to allow users to calculate and visualise the ‘Changes in Phytoplankton and Zooplankton Communities’ (PH1/FW5) indicator, which is used to assess the state of pelagic habitats for the Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR), UK Marine Strategy (UKMS) and Marine Strategy Framework Directive (MSFD).

The app contains all plankton abundance datasets submitted to support the OSPAR Quality Status Report 2023 assessment of Pelagic Habitats and the ongoing data collected for the NC34 Pelagic program. It contains both zooplankton and microphytoplankton abundance datasets. These taxon abundance datasets have been aggregated into plankton lifeform abundance via the Plankton Lifeform Extraction Tool (PLET: https://www.dassh.ac.uk/lifeforms/) and extracted spatially as monthly mean values grouped by their intersection with spatial assessment units (polygons). The OSPAR assessment currently uses the ‘COMP4’ assessment units, and the Water Framework

Directive uses the WER (WFD) assessment units. Both options are available in this tool. Datasets for fixed-point stations are represented as individual points.

Once a dataset has been selected, a relevant lifeform pair is then selected and an appropriate polygon and time period. Two maps and two abundance time periods are displayed together with a graph of the number of samples per months. If there is enough data, the pi tool is also calculated and presented (see the example data blocks for AFBI and the Isle of Man above). The guide to the tool can be found in Appendix 3 and Plankton Lifeform Indicator Tool (PLIT) – Plankton and People.

2.2. Environment Agency

Although much of this data feeds into tasks under other work packages, and on its own is too short for a useful time series the Environment Agency has reviewed the 18 months of zooplankton data so far collected. This has been undertaken in two forms. A descriptive report of taxa found provided by the MBA taxonomists and some exploratory data analysis conducted by EA scientists.

2.2.1. Inshore zooplankton sampling programme (provided by MBA)

Introduction

Since its inception back in July 2022, CPR analysts continue to receive and analyse inshore zooplankton samples taken by Environment Agency (EA) staff from various sampling stations across England as part of the mNCEA programme. At the time of writing, 348 samples have been analysed by the CPR analyst team (this includes samples taken up until the end of 2023). Alongside receipt of each month's samples (per station), accompanying paperwork is included, detailing the meta data of each sample. Measurements such as net volume (m^3), sample position, time of sample and sample depth are included.

This comprehensive sampling programme to date, has given over 18 months' worth of data, helping to fill the gap in zooplankton monitoring from inshore English waters. Furthermore, this component of the mNCEA package complements other sampling efforts undertaken by various other organisations, with the data generated, integrated into other pelagic plankton datasets.

Data

Once each month's samples have been analysed (with 1 in 10 samples reanalysed for quality assurance purposes) and results cross checked, the data is sent to EA personnel in the form of an Excel spreadsheet. At the time of writing, this includes all samples analysed up to the end of November 2023. Additionally, there is one complete year of inshore zooplankton data uploaded to the Marine Environment Data and Information Network (MEDIN) database, whilst all samples up to November 2023 are ready to integrate into the Pelagic Lifeform Extraction Tool (PLET) - this can be implemented once the UK Zooplankton Master List has been reviewed (further detail in section below).

Results

To date, approximately 200 taxa have been recorded across all stations. The most common taxa recorded include small unidentifiable juvenile stages of *Paracalanus*/

Pseudocalanus copepods (known as PPs), unidentified species of *Acartia* spp., *Temora longicornis*, Gastropod larvae, unidentified Polychaete larvae and *Oikopleura* larvaceans (Figure 16).

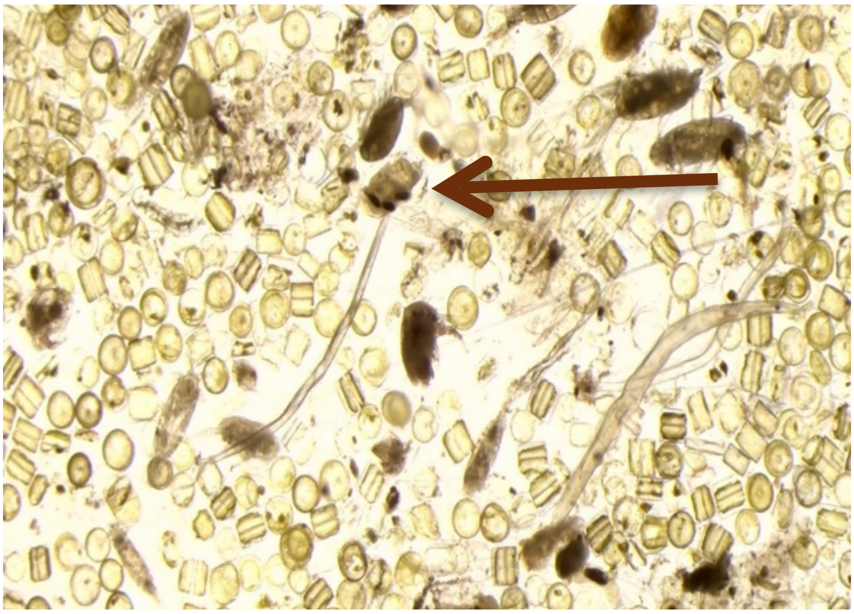


Figure 16. Typical net caught plankton mix, note *Oikopleura* (large head and worm like body). © CPR Survey, Marine Biological Association

The most plankton abundant sample was taken from the sampling station LIV006N at Blackpool (53.83318453, -3.098432663) on the 12 September 2023. On this sample high counts of phytoplankton *Noctiluca* and detritus were noted as well as an abundance of zooplankton.

Centropages hamatus, *Paracalanus*, *PP*s, *Euterpina acutifrons*, Cirripede larvae, Podon, Polychaete larvae, unidentified Hydrozoa and *Oikopleura* were counted in their hundreds, sometimes thousands.

Two samples are also noted as having seahorse fry recorded, (accidentally sampled) found in both Cowes and East Brambles sampling stations (August and October respectively) – EA sampling staff were notified on both occasions.

Figure 17. Seahorse fry © Chris Parkes, Marine Biological Association

UK Zooplankton Master List

Currently there are thirty-two potential additions to the UK Zooplankton Master List which have been identified during this programme, with *Acartia margalefi* a notable copepod not

currently included. Other taxa found by our analysts which are not listed presently have been Cnidarians *Lovenella assimilis* and *Turritopsis polycirrha*. These extra taxa are to be discussed between members of the Pelagic Habitats Expert Group (PHEG), and a decision made on applicable additions to the master list. Once the outcome has been decided then all the data in the PLET template can be uploaded.

Figure 18. Above Left: *Lovenella assimilis*, © CPR Survey, Marine Biological Association, Above Right: *Turritopsis polycirrha*, © CPR Survey, Marine Biological Association

Non-natives

Originally native to the Indo-Pacific area, the copepod *Pseudodiaptomus marinus* has spread across European wates over the last decade with sightings around English waters increasing; in fact, it is thought to be the species probably spreading into new areas at the fastest pace. In these samples *P. marinus* has been frequently recorded, found in multiple samples taken around the Solent, Kent, Essex and Lincolnshire areas in late autumn and early spring months. The highest number recorded was found from a sample taken at Virley Channel (tidal channel in the estuary of River Blackwater, Essex) during early October 2023 with 471.34 counted per m^3.

Sample data for this copepod has been disseminated to the ICES European Observatory of the Non-Indigenous Calanoid Copepod *Pseudodiaptomus marinus* (EUROBUS) working group who are investigating this species more closely.

The most diverse copepod genus found on samples to date has been *Acartia* with five species recorded. Two non-native species of *Acartia* have been identified – *Acartia tonsa* and *Acartia margalefi*, both being more suited to warmer Mediterranean waters. *A. margalefi* is smaller than other species of *Acartia* and is normally found in estuarine environments, though is thought to tolerate a range of temperatures. On these samples it was found in spring, summer, and autumn months from stations in the Solent and Virley Channel, Essex – the highest number recorded at Fawley in March 2023 with 92.84 organisms per m^3. *A. tonsa* was found in autumn months from stations in Kent, Mersey Estuary, and Inner Bristol Channel, the most frequent being 8.49 per m^3 at the Virley Channel station.

Tortanus discaudatus, a neritic copepod previously recorded from North-West Atlantic samples was first recorded in the 1960s by the CPR Survey on a tow west of Ireland and analysts have since been observing their presence in North Sea CPR samples over recent years. So far, the team have recorded *T. discaudatus* in five samples, all from the Yorkshire and Northumberland coastline.

Two non-native jellyfish have also been recorded – *Lovenella assimilis* and *Turritopsis polycirrha. L. assimilis* was previously thought to prefer tropical to temperature areas, but was first documented in the Bay of Biscay, then the English Channel a couple decades ago. This species has been recorded around the Solent stations from August – November 2023 and a couple of South Kent stations in October and November 2023 (see images above).

2.2.2. Exploratory statistics

In the year 1 report our limited data set (about 7 months of data) had identified 152 taxa from a 'potential' list of 231 held by MBA analysts. Now with about 18 months of data this has increased to 192 taxa. In 2022, the total estimate count was almost 200,000 organisms which has now risen to a little over 794,000 organisms. The increase in diversity and numbers reflects the extra months and seasons sampled. Only 10 taxa represent over 75% of that total, while only 3 taxa contribute more than 10%. These are dominated by the copepods *Acartia, Temora, Centropages, Paracalanus, Euterpina, and* Calanoida; Cirripedia (barnacle) and gastropod larvae; the Appendicularia *Oikopleura*. (Table 9).

	Total counts per m3 by Region								
Taxon	**Anglian**	**NEast**	**NWest**	**Southern**	**SWest**	**Thames**	**Grand Total**	**% of total**	**Cumulative Percentage**
Acartia	24832.2	55022.8	5361.14	37434.57	11026.02	2567.45	**136244.3**	17.16%	17.16%
Calanoida	25268.9	30819.8	13413.1	24594.30	11222.69	1949.68	**107268.5**	13.51%	30.66%
Cirripedia	15220.1	13984.6	25004.3	20635.99	5804.22	5195.22	**85844.53**	10.81%	41.47%
Temora longicornis	9984.99	13295.05	8154.35	22923.77	4268.26	446.01	**59072.44**	7.44%	48.91%
Centropages hamatus	21348.68	2463.81	12854.58	13923.53	326.23	856.10	**51772.94**	6.52%	55.43%
Gastropoda	7011.13	3808.41	4876.14	19952.13	3035.74	1265.35	**39948.89**	5.03%	60.46%
Oikopleura	6543.89	4753.30	15814.0	6815.87	2294.12	1423.79	**37645.04**	4.74%	65.20%
Acartia (Acartiura) clausi	6386.01	12419.17	487.14	11624.33	4100.40	1096.84	**36113.89**	4.55%	69.75%
Euterpina acutifrons	5083.58	10.40	6142.07	7899.08	3311.74	5896.24	**28343.11**	3.57%	73.32%
Paracalanus	3165.79	5781.87	3586.42	5028.80	5879.01	426.92	**23868.80**	3.01%	76.32%

Table 9. Top 10 taxa contribution to 75% of the total organisms counted.

Considering the most frequently encountered taxa groups, across all water bodies and sampling occasions, only 6 taxa have a prevalence of over 60% at a national scale (Table 10).

	Frequency of occurrence in Region							
Taxon	**Anglian**	**NEast**	**NWest**	**Southern**	**SWest**	**Thames**	**Grand Total**	**Percentage of Max possible**
Calanoida	71	66	50	98	40	11	**336**	71.79%
Acartia	72	66	40	97	35	11	**321**	68.59%
Gastropoda	69	63	44	96	38	11	**321**	68.59%

Polychaeta	63	51	47	91	36	11	**299**	63.89%
Temora longicornis	61	65	40	85	37	10	**298**	63.68%
Oikopleura	50	50	49	95	36	11	**291**	62.18%

Table 10. Most prevalent taxa (occurring in over 60% of samples)

The percentage taxa from the full list tended to be in the order of 50-70% (Table 11). The exception is the Thames, which is a very small area with relatively few sites compared to the other regions.

	Anglian	NEast	NWest	Southern	SWest	Thames
Count	118	132	102	132	105	72
Percentage	61%	69%	53%	69%	55%	38%

Table 11. Prevalence per EA region

If we look at the top 25 taxa in each group, we see that 18 taxa occur in both groups and can be considered the "most common" in these samples (highlighted in bold in Table 12).

	Abundance		Prevalence	
Taxon	**Percentage of total**	**Cumulative Percentage**	**Taxon**	**Percentage of Max possible**
Acartia	17.16%	17.16%	**Calanoida**	71.79%
Calanoida	13.51%	30.66%	**Acartia**	68.59%
Cirripedia	10.81%	41.47%	**Gastropoda**	68.59%
Temora longicornis	7.44%	48.91%	**Polychaeta**	63.89%
Centropages hamatus	6.52%	55.43%	**Temora longicornis**	63.68%
Gastropoda	5.03%	60.46%	**Oikopleura**	62.18%
Oikopleura	4.74%	65.20%	**Cirripedia**	59.19%
Acartia (Acartiura) clausi	4.55%	69.75%	**Paracalanus**	56.84%
Euterpina acutifrons	3.57%	73.32%	**Bryozoa**	56.41%
Paracalanus	3.01%	76.32%	**Centropages hamatus**	55.98%
Polychaeta	2.62%	78.94%	**Brachyura**	55.56%
Oithona	2.42%	81.37%	**Acartia (Acartiura) clausi**	54.70%
Bivalvia	1.92%	83.29%	Bivalvia	49.36%
Podon	1.81%	85.10%	**Euterpina acutifrons**	43.80%
Bryozoa	1.68%	86.78%	Caridea	42.74%
Pisidia longicornis	1.09%	87.88%	Pleurobrachia pileus	40.81%
Acartia (Acanthacartia) bifilosa	1.09%	88.96%	**Oithona**	39.74%
Evadne nordmanni	1.03%	89.99%	Hydrozoa	39.32%
Pseudocalanus	0.98%	90.97%	**Parasagitta setosa**	37.82%
Isias clavipes	0.70%	91.67%	**Pseudocalanus**	37.61%
Copepoda	0.69%	92.36%	Cyclopoida	36.11%

Brachyura	0.67%	93.03%	**Copepoda**	35.47%
Acartia (Acartiura) discaudata	0.57%	93.60%	Chaetognatha	31.84%
Foraminifera	0.49%	94.09%	**Pisidia longicornis**	30.56%
Parasagitta setosa	0.47%	94.55%	Foraminifera	28.85%

Table 12. Comparison of most abundant and most prevalent taxa.

These results are similar, but not identical to, last year. The dissimilarity is mainly due to the large number of samples and temporal coverage achieved.

A taxonomic tree can be constructed from the present and absence data, and this suggests we have a broad spectrum of inshore taxa and all the key expected groups (Figure 19).

Figure 19. Taxonomic tree constructed from Environment Agency inshore samples.

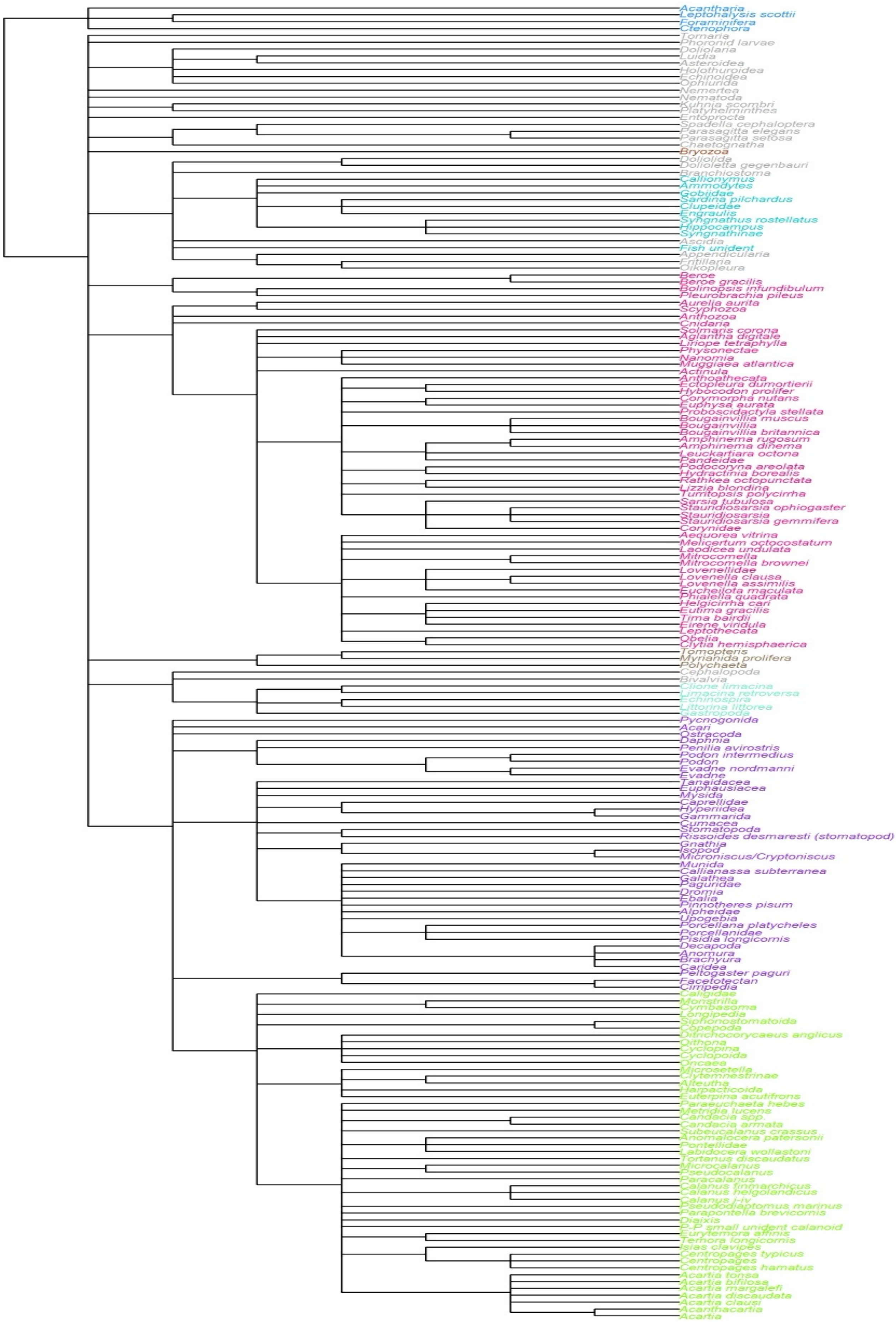
2.2.3. Zooplankton statistical models

We have investigated a model-based approach to explore underlying trends in the zooplankton abundance data and how these relate to water quality parameters. The underlying zooplankton data have been gathered since mid-2022 as part of the Environment Agency's monitoring programme. As these samples were gathered at the same time and place as the Environment Agency's routine water quality monitoring program, each zooplankton sample has a water quality sample associated with it.

Water quality data were extracted from the Environment Agency's Water Information Management System (WIMS) data base. Zooplankton density data were matched to the WIMS sample data through cross-referencing with the associated PRN value, which provide a unique reference for each sampling event. The following water quality parameters were extracted from the WIMS data:

- Ammonium (NH_4) concentration (mg/l)
- Dissolved inorganic nitrogen (DIN) concentration (mg/l)
- Orthophosphate (PO_4) concentration (mg/l)
- Salinity (ppt)
- Chlorophyll *a* concentration (µg/l)
- Water depth (m)
- Water temperature (°C)

This resulted in a sample-by-taxon table and a sample-by-water quality table. These data were analysed using a generalised linear latent variable model (using the R package "gllvm", Niku et al., 2019).



Prior to analysis, water quality parameter values recorded at lower than the analytical equipment's limit of detection (LOD) were replaced with values of 0.5*LOD. For example, values recorded as "<0.1 mg/l" were recorded as 0.5*0.1 = 0.05 mg/l. Furthermore, as the GLLVM software does not allow for missing values in the analyses, missing values in the

WIMS data were replaced with the mean value for the given variable. Using the mean value minimises the effect that these ‘missing’ values have on the model calculations.

One reason for the shift towards statistical model-based approaches is that distance-based approaches do not account for the strong mean-variance relationship common to ecological data. That is, the models underlying these traditional approaches are based on the assumption of homogenous variability, regardless of mean taxon abundance. However, this is not the case, with variance levels clearly showing a strong positive relationship with mean taxon abundances, as represented in Figure 20.

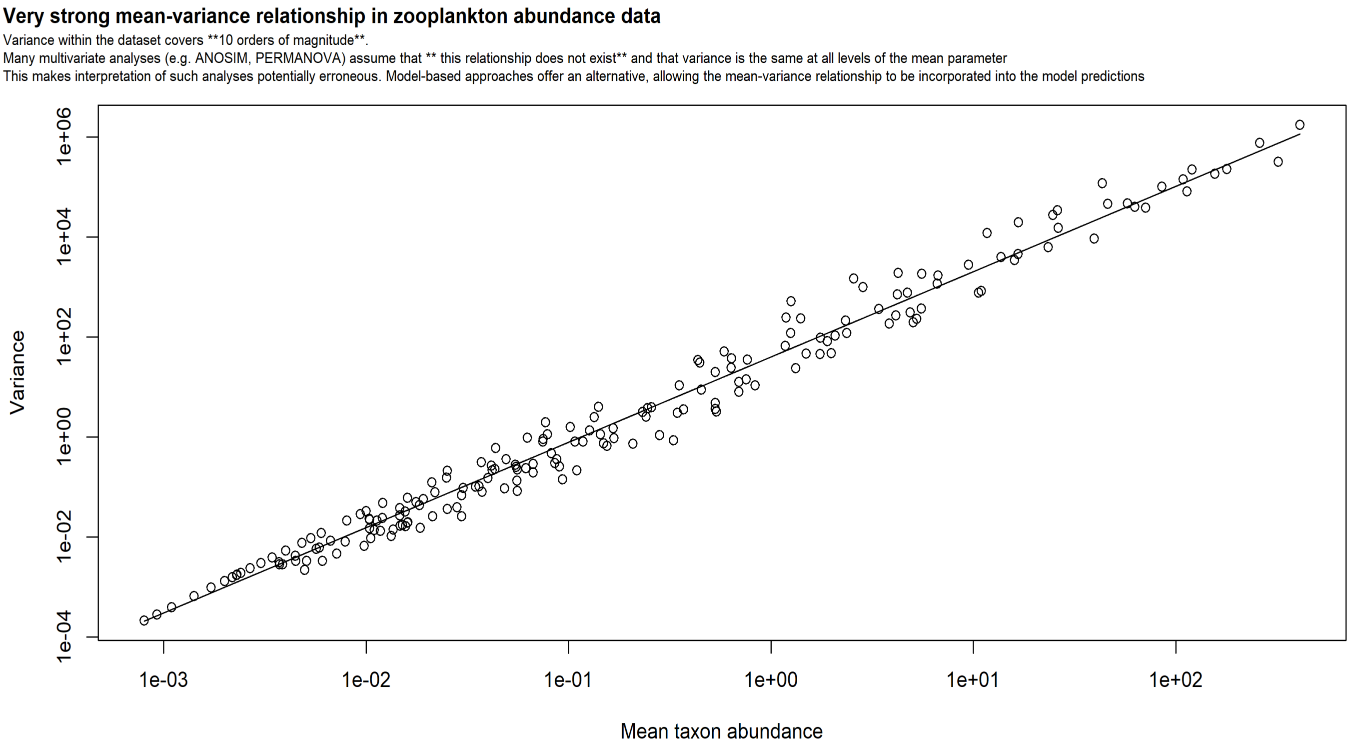
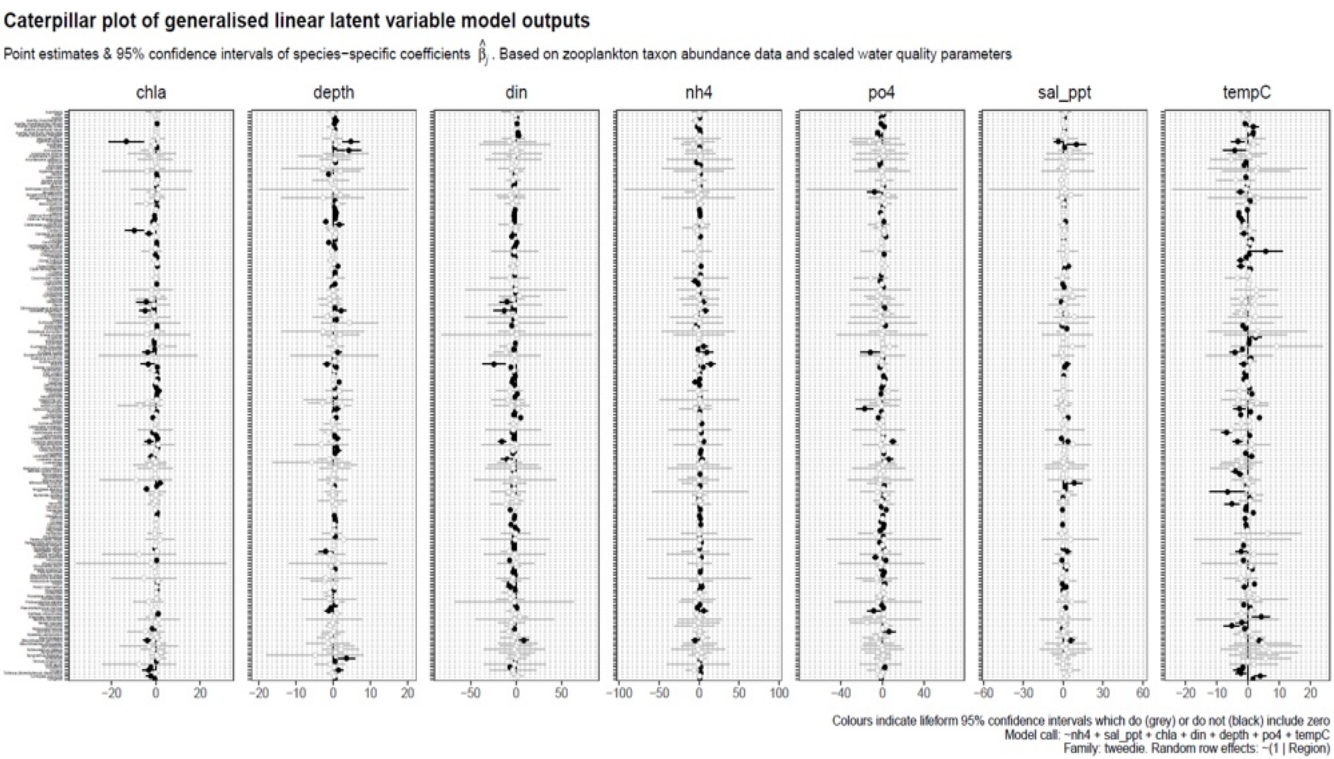


Figure 20. Zooplankton data showing strong relationship with variance.

To graphically display the gllvm model outputs, the coefficients of the first latent variable were extracted for each taxon or lifeform. This comprised a point estimate and a 95% confidence interval. These were then plotted, with taxon/lifeforms whose confidence intervals did not cross over zero highlighted.



We have looked at this both at the taxa level, and the more ecologically informative lifeform level. These have been expressed as caterpillar plots (Figures 21 and 22).

Figure 21. Caterpillar plot of taxa against environmental variables.

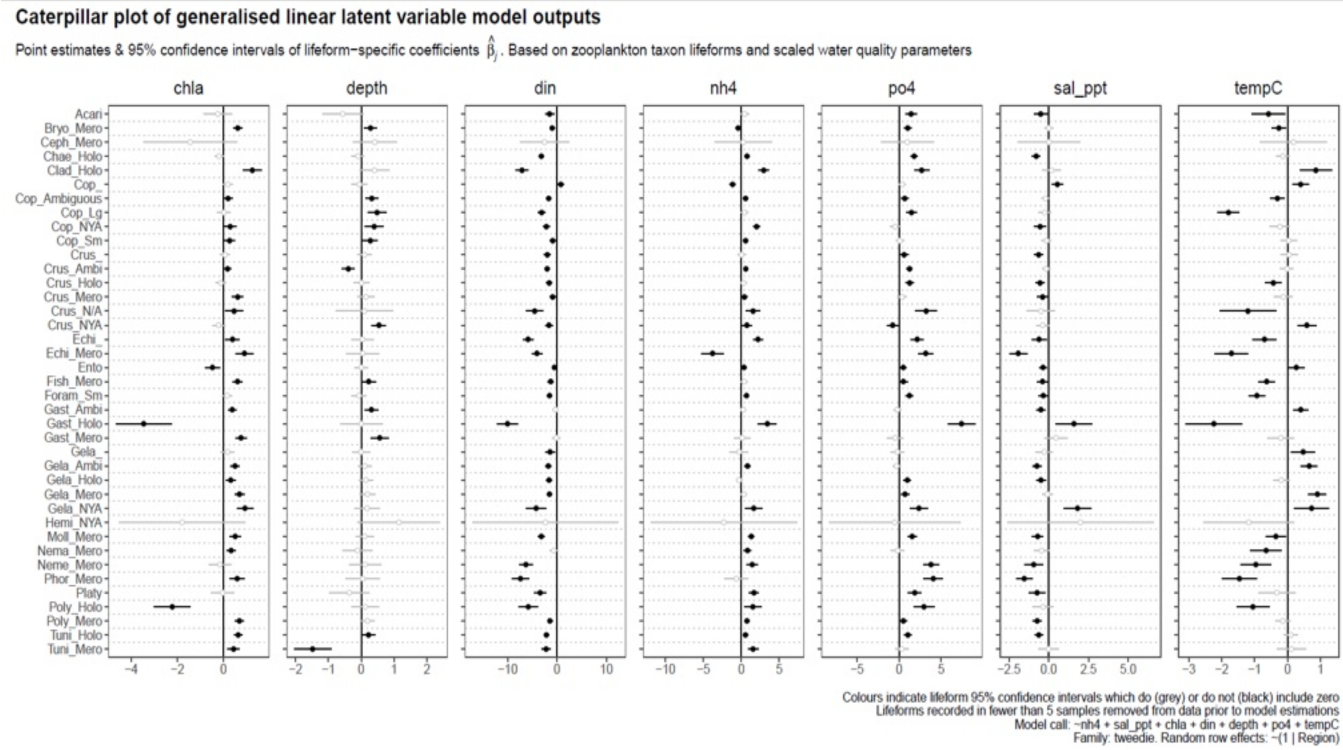


Figure 22. Caterpillar plot of lifeforms against environmental variables.

For more details and larger graphics see appendix 4.

2.3. Cefas

There was little variation in the mean abundance of zooplankton between sites (indicated by the bold black line in the middle of each box, Figure 23). The ‘East of Warp Site’ located in the Thames estuary had the highest mean species abundance across all samples and ‘Liverpool Bay SmartBuoy’, located in the Mersey estuary, had the lowest (Figure 23).

The highest individual species abundances (indicated by the individual black dots) were found at the Thames sites and the lowest found in the Mersey site (Figure 23). The Thames sites had very similar ranges in terms of individual species abundances (noted by difference between the highest and lowest black dots for each site on Figure 24).

Monthly fluctuations in both estuaries mean abundance (indicated by the bold black line in the middle of each box, Figure 24) and individual species abundance (indicated by the individual dots in Figure 24x between estuaries. Both estuaries had their highest mean and individual species abundance in summer months with lower mean species abundance in winter months.

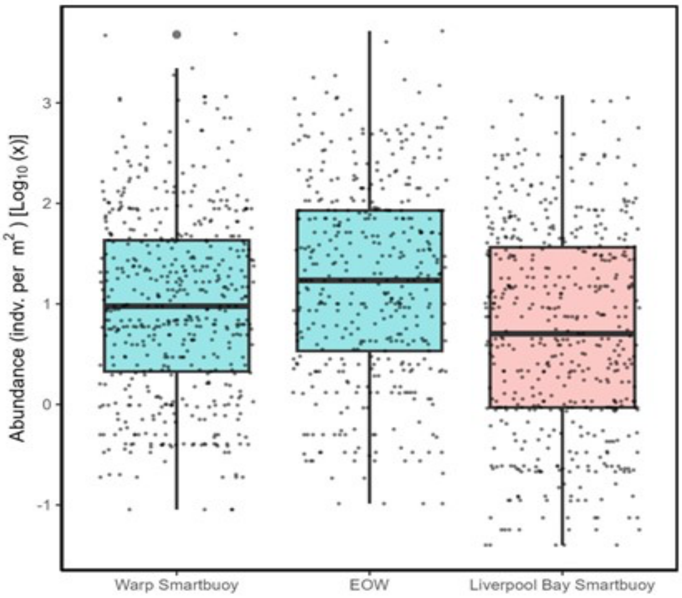


Figure 23. Boxplot of zooplankton abundance for all three sites for all species and all samples. Each black dot represents the individual abundance of a single species in a single sample. Colour indicates estuary where: Blue = Thames, Pink = Mersey.

There were some individual species that were continued to be found in high abundances in winter months at both estuaries. I off in the autumn and winter (although some individual taxa still exhibit high numbers even in January). In general, the interquartile range (indicated by the top and bottom of each box in Figure 24) was reduced as the year progressed.

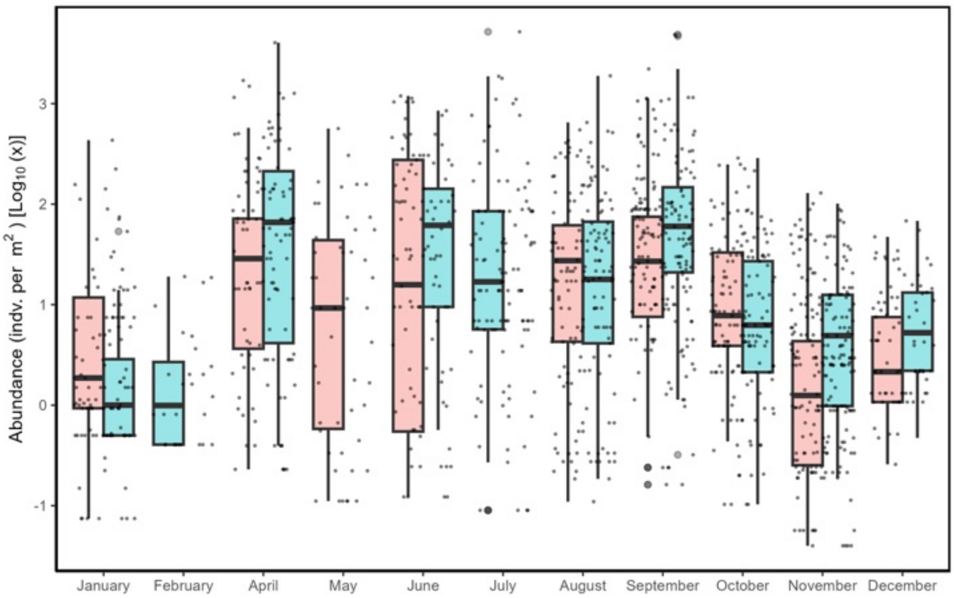


Figure 24. Boxplot of zooplankton abundance for each estuary per month. Each black dot represents the individual abundance of a single species in a single sample. Colour indicates estuary where: Blue = Thames, Pink = Mersey.

The 10 most abundance zooplankton species per site is shown in Table 13. There is some similarity between the sites and estuaries in terms of the most abundance taxa. For example, calanoid copepods (Calanoida C1-6), Cladocera and Temora *spp*. among others are found at all three sites.

Warp Smartbouy	**EOW**	**Liverpool Bay Smartbuoy**

Taxa	Mean Abundance (indv per m^3)	Taxa	Mean Abundance (indv per m^3)	Taxa	Mean Abundance (indv per m^3)
Cladocera	1951.24	Cladocera	1212.59	Evadne	639.97
Euterpina acutifrons	370.66	Temora	534.23	Appendiculari a	213.66
Sabellaria	217.05	Cirripedia	263.16	Cirripedia	162.18
Echinodermata	172.42	Temora longicornis	225.26	Oithona	136.28
Cirripedia	127.83	Foraminifera	217.53	Echinodermata	122.42
Appendicularia	119.86	Acartia	205.26	Calanoida C1-6	119.95
Lamellibranchiata	118.53	Calanoida C1-6	196.03	Cladocera	110.72
Calanoida C1-6	115.07	Centropages	182.24	Paracalanus	104.99
Temora	107.47	Appendicularia	181.33	Copepoda C1-6	88.32
Acartia	97.69	Euterpina acutifrons	159.82	Temora	72.03

Table 13. Most abundance taxa (mean abundance across all months) for each of the three sites. EOW = East or warp. n indicates the commonly found taxa.

2.4. MBA and mNCEA funded CPR routes

May 2022 saw several routine CPR English routes reinstated under the mNCEA programme. These routes are from the NE Atlantic to Southwest Approaches (BB, BC, BD), Dogger Bank (KC) and Southern North Sea (HE). To date, analysis has been completed for all reinstated routes from May – December 2022, analysis of which has followed the normal CPR methodology, with every other sample being analysed. The CPR team are now working on analysis of 2023 samples.

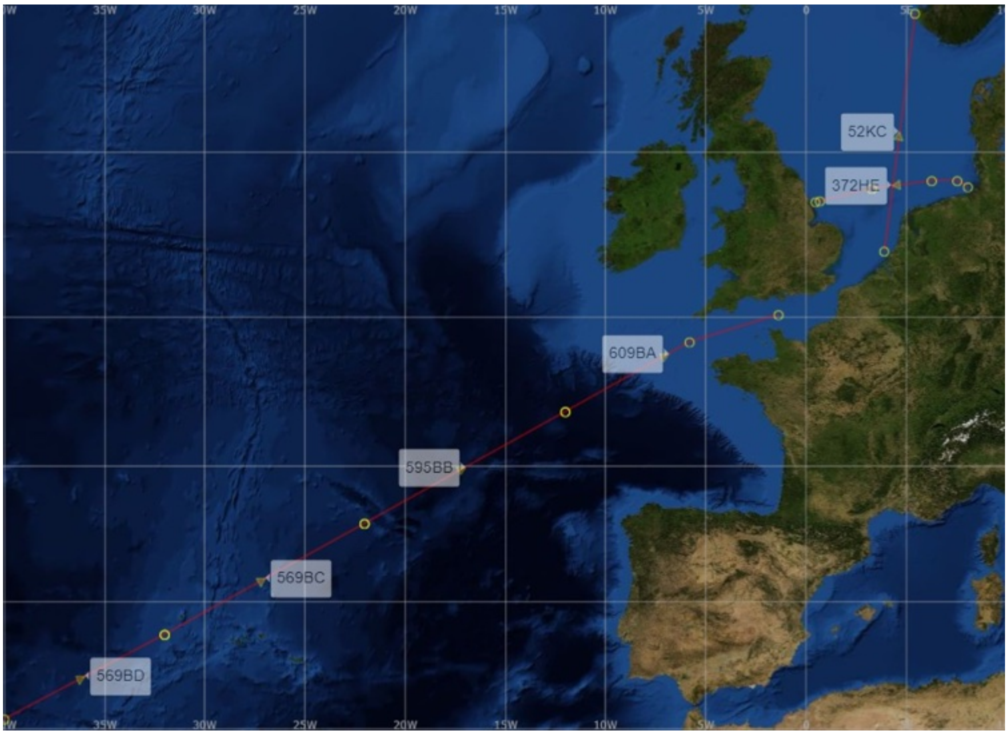


Figure 25. Continuous Plankton Recorder (CPR) tows: 569BD, 569BC, 595BB, 609BA, 52KC, 372HE towed regularly over 2022/23.

As well as analysis of plankton data, every single CPR sample gets assigned as value for 'greenness' before it is cut and distributed to the analysis team. This is known as Phytoplankton Colour Index (PCI) and gives a semi-quantitative representation of the total phytoplankton biomass. Furthermore, it includes the organisms that are too fragile to survive the sampling process. PCI can also be used as a comparison against satellite data and in conjunction with historical data.

The colour estimates are related to the numbered divisions of the graduated silk and subsequently to the samples when the record has been cut. The colour categories are shown in the table below.

Category	Recorded
No Colour	NIL (0)
Very pale green	VPG (1)
Pale green	PG (2)
Green	G (3)

Figure 26. Left: CPR trained staff undertaking PCI process. Above: PCI colour categories

PCI data has been collected on these reinstated routes and available once each tow has been processed – this currently includes all samples towed up to and including, October 2023.

3. Data management plans

Raw data is obtained from samples collected by the monitoring institutions:

- Agri-Food and Biosciences Institute (AFBI)
- Centre for Environment, Fisheries and Aquaculture Science (Cefas)
- Environment Agency (EA)
- Isle of Man Government (IoM)
- Marine Biological Association (MBA)
- Marine Scotland Science (MSS), now The Scottish Ministers acting through Marine Directorate (SEDD) rather than MSS.
- Newcastle University (Dove Marine Laboratory)
- Plymouth Marine Laboratory (PML)
- Scottish Association for Marine Science (SAMS)

Each of these institutes collect samples under their various sample programs and manged through their proprietary sampling management programs. The collected samples are identified either via in house or external laboratories. Returned taxa identifications and counts are stored on each institute's systems and managed through their own data management plans. Most organisations will follow Q-FAIR and Medin standards and may be their own Data Archive Centre (DAC).

The plankton data collected will be put on Plankton Lifeform Extraction tool (PLET) which is part of DASSH, under MEDIN. Each data set has a data object identifier (DOI) to enable discoverability. Data upload to PLET is an ongoing process. Table 15 gives an indication of data available on PLET.

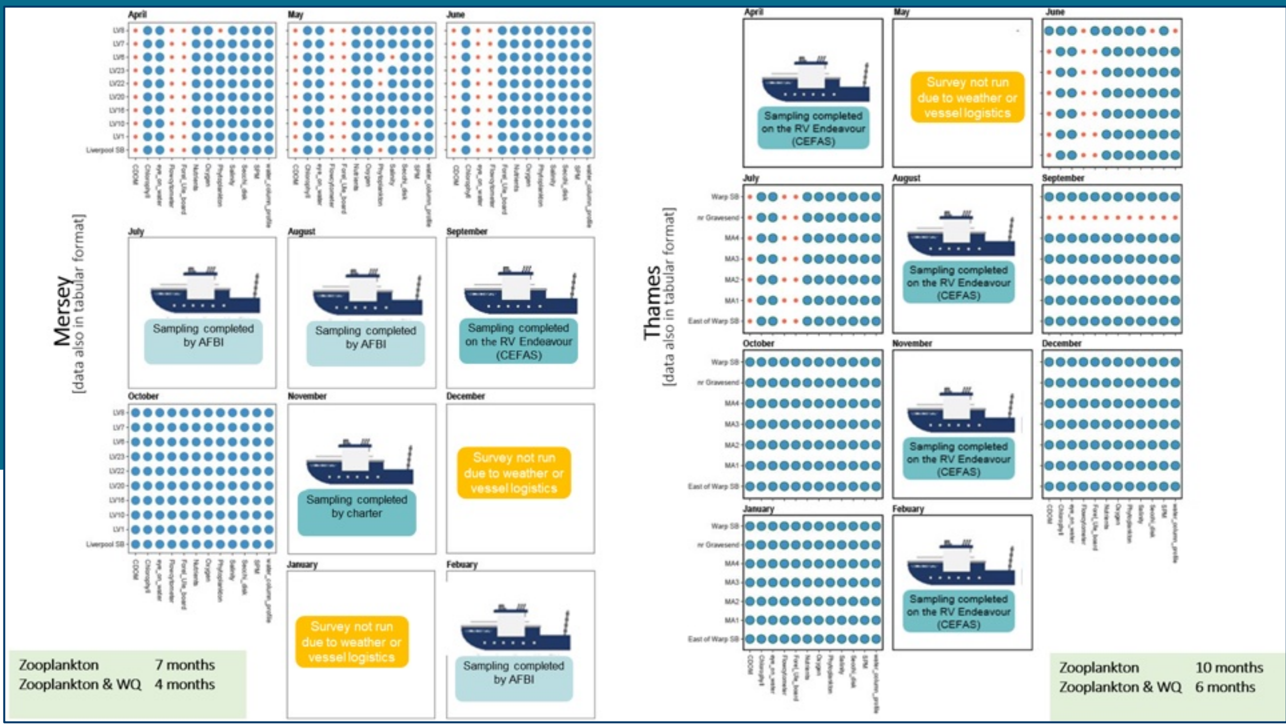
Data is extracted either from PLET or the institution's data archived, combine with partner institutions data and other data sets, and manipulated to produce a *derived* data set for the various workstreams within the other work packages. These derived data are stored locally but may be provided as supplementary information in papers and report.

Data management plans are being developed for the project, with the over-arching goal to ensure this is open source.

Institute	Dataset	Unique sites	Unique samples	Earliest year on PLET	Latest year on PLET
UK-Agri-Food and Biosciences Institute (AFBI)	Western Irish Sea Long term monitoring	1	54	2015	2021
UK-Centre for Environment, Fisheries and Aquaculture Science (Cefas)	Cefas SmartBuoy Marine Observational Network - UK Waters Phytoplankton Data 2001-2019	7	901	2001	2019
	Cefas West Gabbard zooplankton abundance time series monitoring using ZooScan from 2016 to 2020	1	38	2016	2020
UK-Environment Agency (EA)	EA CHL 2000-2020	27,011	40,531	2000	2020
	EA PHYTO 2000-2020	21,705	28,668	2000	2020
UK-Isle of Man Government (IOM)	Cypris Station Phytoplankton Abundance	5	817	1995	2012
UK-Marine Biological Association (MBA)	Continuous Plankton Recorder	128,577	133,70[illegible]	1960	2019
UK-Marine Scotland (MS)	MSS Loch Ewe Phytoplankton	2	2,605	4002	4040
	MSS Loch Ewe zooplankton	1	820	2002	2017
	MSS Scalloway Phytoplankton dataset	1	703	2000	2018
	MSS Scapa Phytoplankton dataset	1	957	2000	2020
	MSS Stonehaven Chlorophyll data	1	1,018	1997	2020
	MSS Stonehaven Phytoplankton	1	875	2000	2020
	MSS Stonehaven zooplankton	1	962	1999	2020
UK-Natural Resources Wales (NRW)	NRW WFD Phytoplankton classification data 2007 - 2019	119	8,167	2007	2019
UK-Newcastle University (NU)	Dove Time Series Ichthyoplankton zoo	1	215	1972	2008
	Dove Time Series WP2 Microscope and FlowCam	1	401	1969	2008
	Newcastle University Dove Time Series WP2 and Horizontal WP3	1	293	1969	1996
	Newcastle University Dove Time Series WP2 FlowCam zoo	1	108	1997	2008
	Newcastle University/Cefas Dove Time Series P200	1	281	1971	2002
UK-Plymouth Marine Laboratory (PML)	PML_L4 chl a	1	1,713	1992	2019
	PML_L4 phytoplankton	1	1,183	1992	2020
	PML_L4 zooplankton	1	1,451	1988	2020
UK-Scottish Association for Marine Science (SAMS)	SAMS-LPO-Phyto-Dec2021	7	1,295	1970	2017

Table 15. Datasets currently in PLET together with dataset metrics. Note this does not reflect the latest data collection and discoveries.

1. Infographic of increased sampling in Thames and Mersey
2. Review of Novel Methods
3. Plankton Lifeform Indicator Tool
4. Zooplankton statistical models



Larger version of Figure 4. Success of increased sampling in the Thames and Mersey supported by EA survey boats, Cefas RV Endeavour and AFBI research vessels. The blue and orange dots represent the water quality and plankton samples that are taken on the EA survey boats, with orange dots in the earlier months highlighting what samples were not collected, and blue dots where the sample had been successfully taken. The move from orange to blue dots in later months showing the success of adding new parameters to the EA surveys including CDOM, flow cytometry and ocean colour.

The report on novel methods embedded below focusses on imaging (e.g. Flow Cam and Flow cytometry) and basic molecular methods.

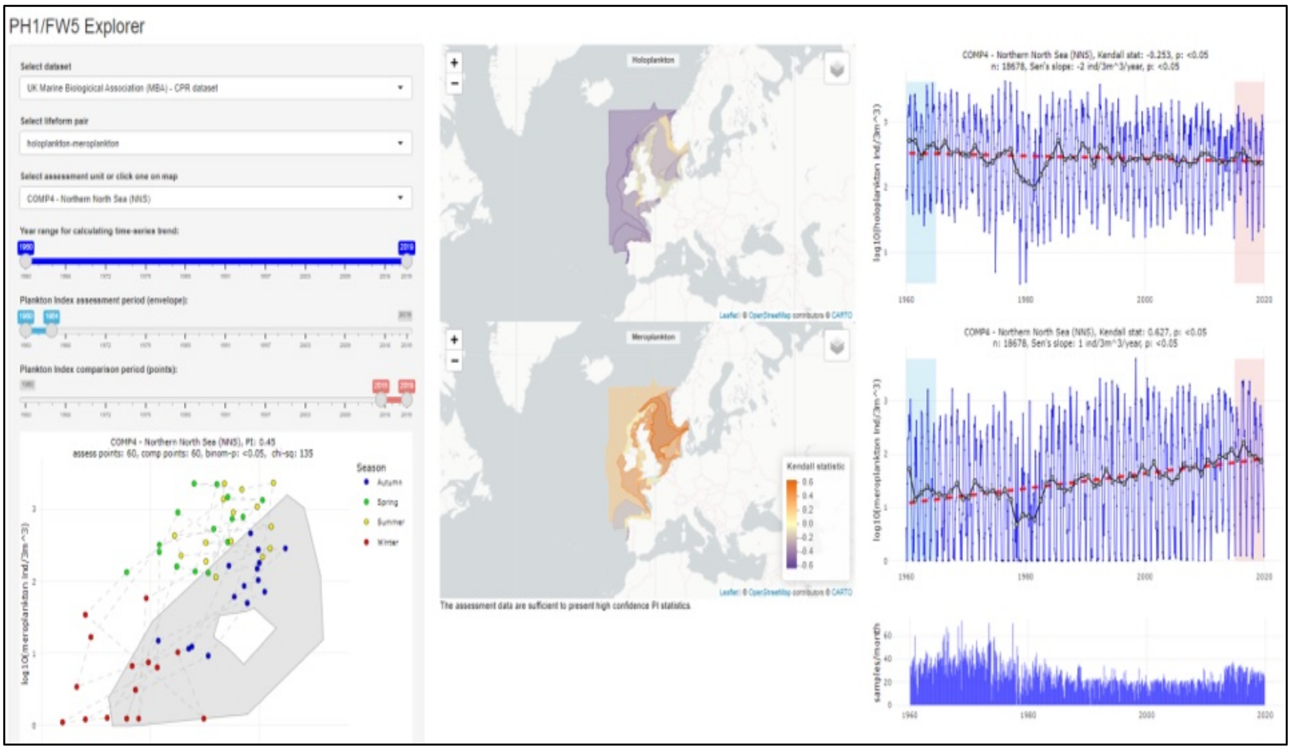
It concludes:

- This report presents a short overview of zooplankton monitoring method currently in use.
- Zooplankton methods presented include imaging, flowcytometry and DNA.
- An overall UK monitoring plan should include be a hybrid using, as appropriate, all types of methods, with a mix of long term established methods collected alongside high frequency imaging and molecular measurements.
- The mix of methods will depend on the monitoring program, and the assessment area and the type of data that is required.
- Long term established plankton monitoring sites should be protected, and any changes in terms of methodology or sampling equipment needs to be carefully considered in terms of protecting the long-term data legacy.
- Ongoing national monitoring programs need to consider how to embed the newer methods, and consider long term impacts on data, costs, and efficiency.

mNCEA_format_Report on novel method

Appendix 3. Plankton Lifeform Indicator Tool (PLIT)

The PLIT is described on the People and Plankton website (Plankton Lifeform Indicator Tool (PLIT) – Plankton and People) and is reproduced below.



Use the Plankton Lifeform Indicator Tool (PLIT)

Introduction

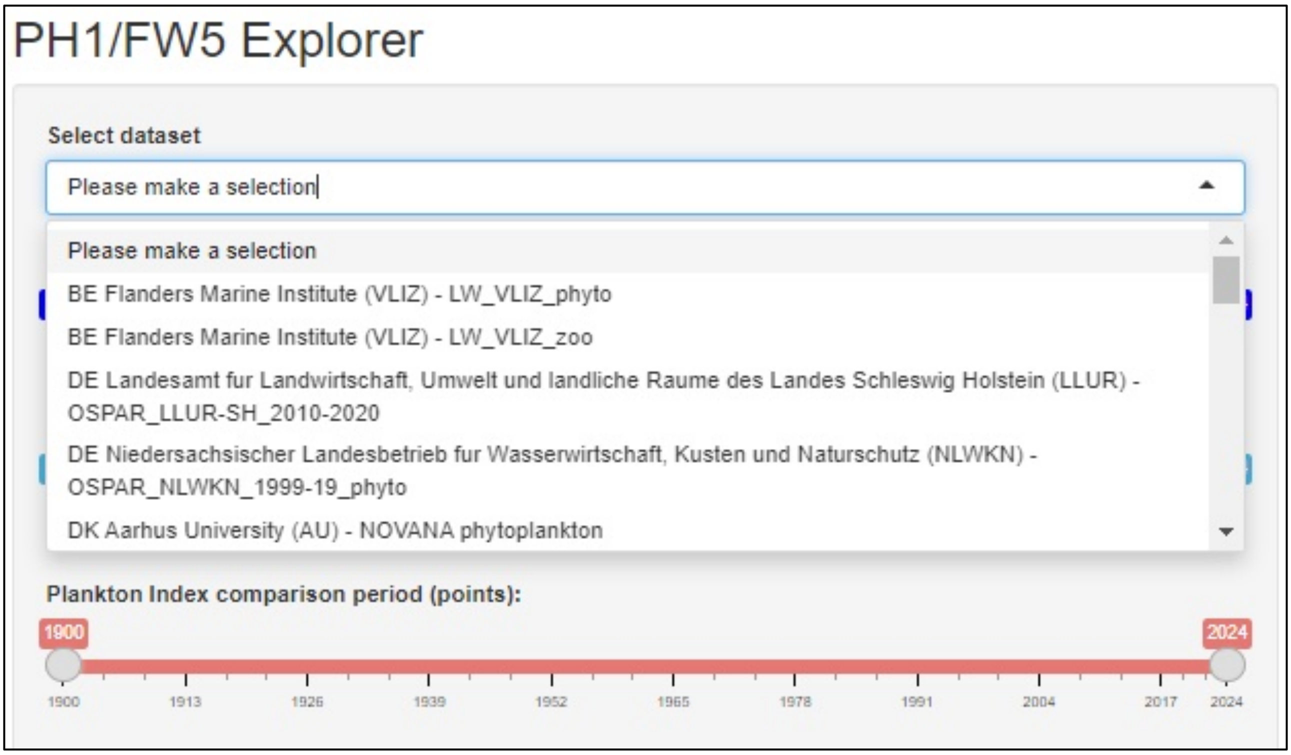
This web-based tool allows users to calculate and visualise the Changes in Phytoplankton and Zooplankton Communities (PH1/FW5) indicator, which is used to assess the state of pelagic habitats for OSPAR, UK Marine Strategy (UKMS) and Marine Strategy Framework Directive (MSFD). For information on how the PH1/FW5 indicator is calculated, refer to: https://www.ospar.org/documents?v=39001.

This indicator evaluates covariance in the abundance of plankton functional groups or "lifeforms" through time. The lifeforms in each pair are considered to be ecologically linked in some way and their co-variance through time reveals information about changes in the pelagic environment. Indicators based on functional groups have been proven relevant for the description of the community’s structure and biodiversity and are more easily inter-compared than other indicators based on taxonomy.

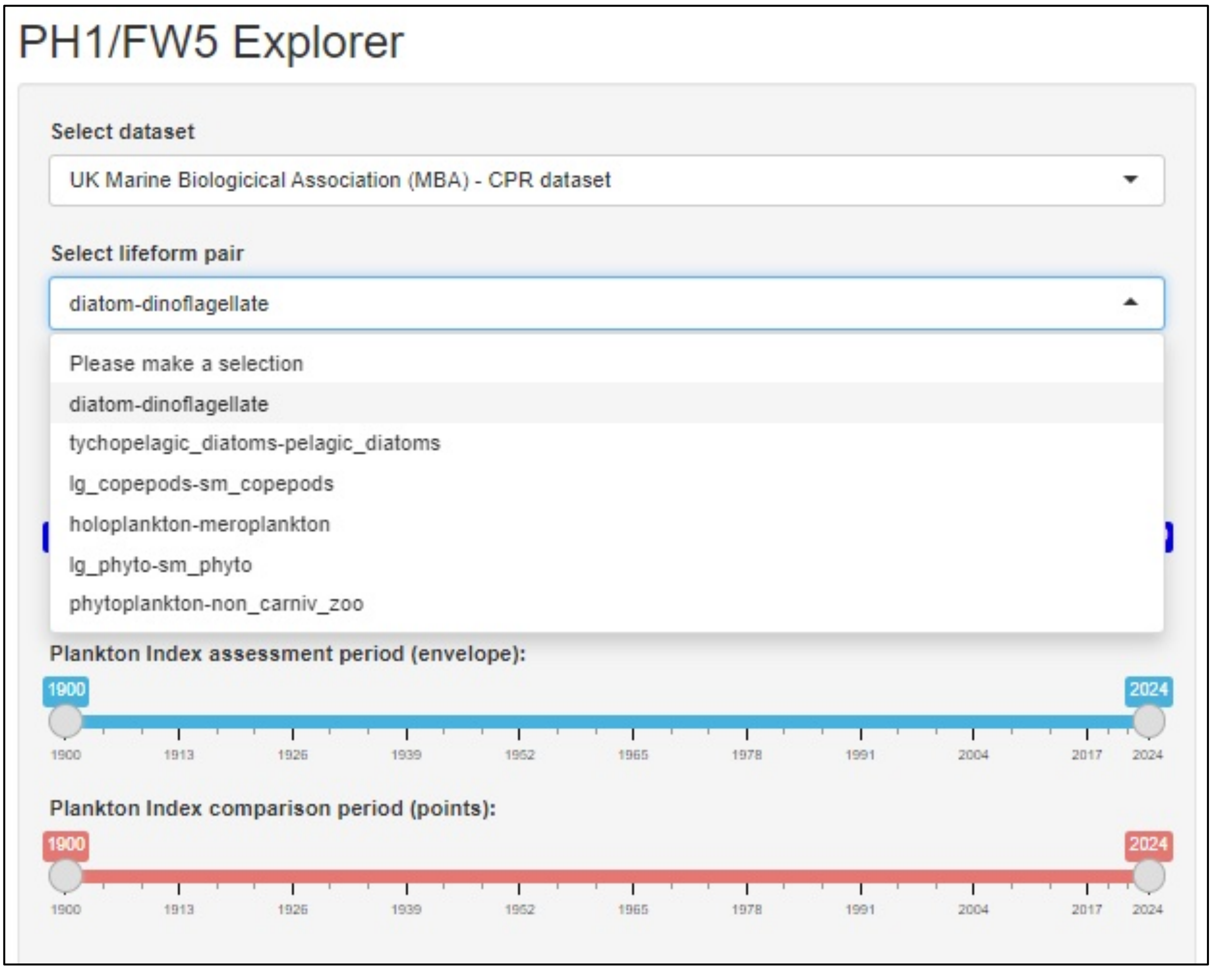
The app contains all plankton abundance datasets submitted to support the OSPAR Quality Status Report 2023 assessment of Pelagic Habitats. It contains both zooplankton and microphytoplankton abundance data from 9 countries that are part of the OSPAR Convention, also known as "OSPAR contracting parties". These taxon abundance datasets have been aggregated into plankton lifeform abundance via the Plankton Lifeform Extraction Tool (PLET: https://www.dassh.ac.uk/lifeforms/) and extracted spatially as monthly mean values grouped by their intersection with spatial assessment units (polygons). The OSPAR assessment currently uses the COMP4 assessment units, and the Water Framework Directive uses the WFD assessment units. Both options are available in this tool. Datasets for fixed-point stations are represented as individual points.

Instructions

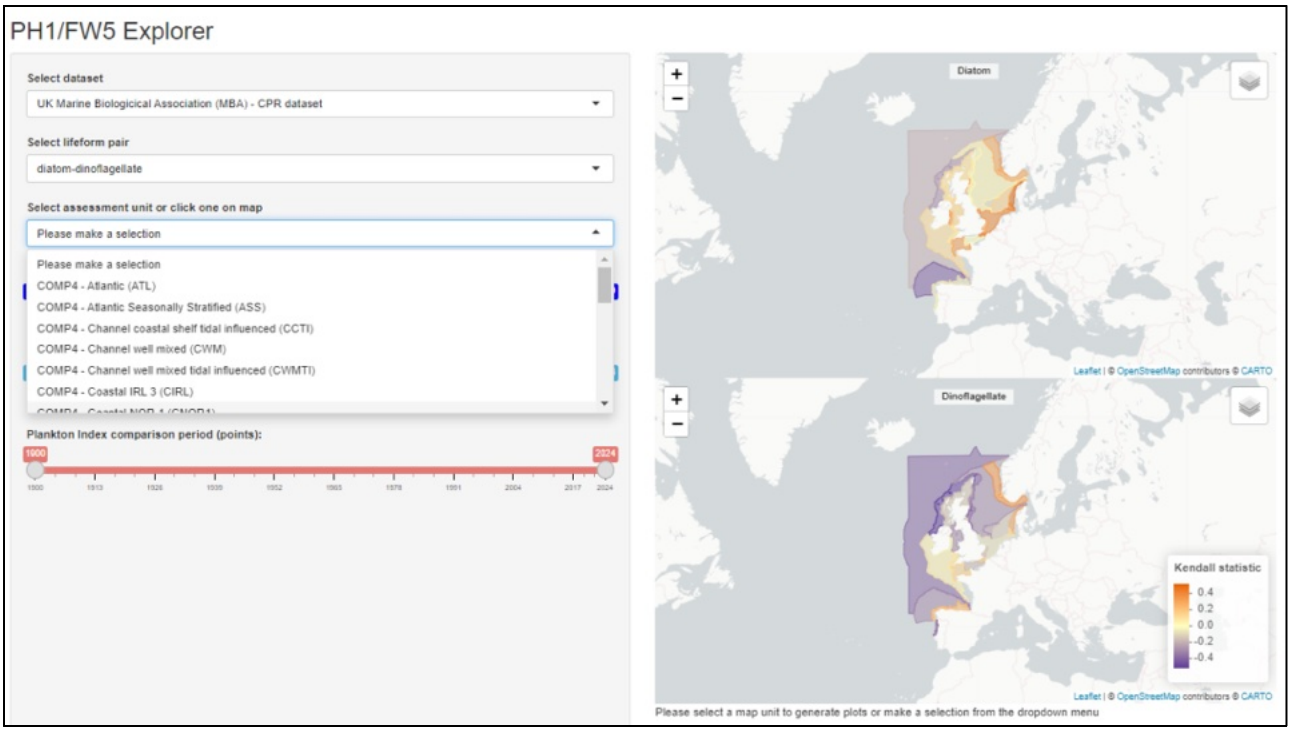
To use the tool, users must first select a plankton abundance dataset from the **Select dataset** drop-down menu. Datasets are organised by OSPAR contracting party in alphabetical order (BE: Belgium, DE: Germany, DK: Denmark, ES: Spain, FR: France, NL: Netherlands, PT: Portugal, SE: Sweden, UK: United Kingdom).



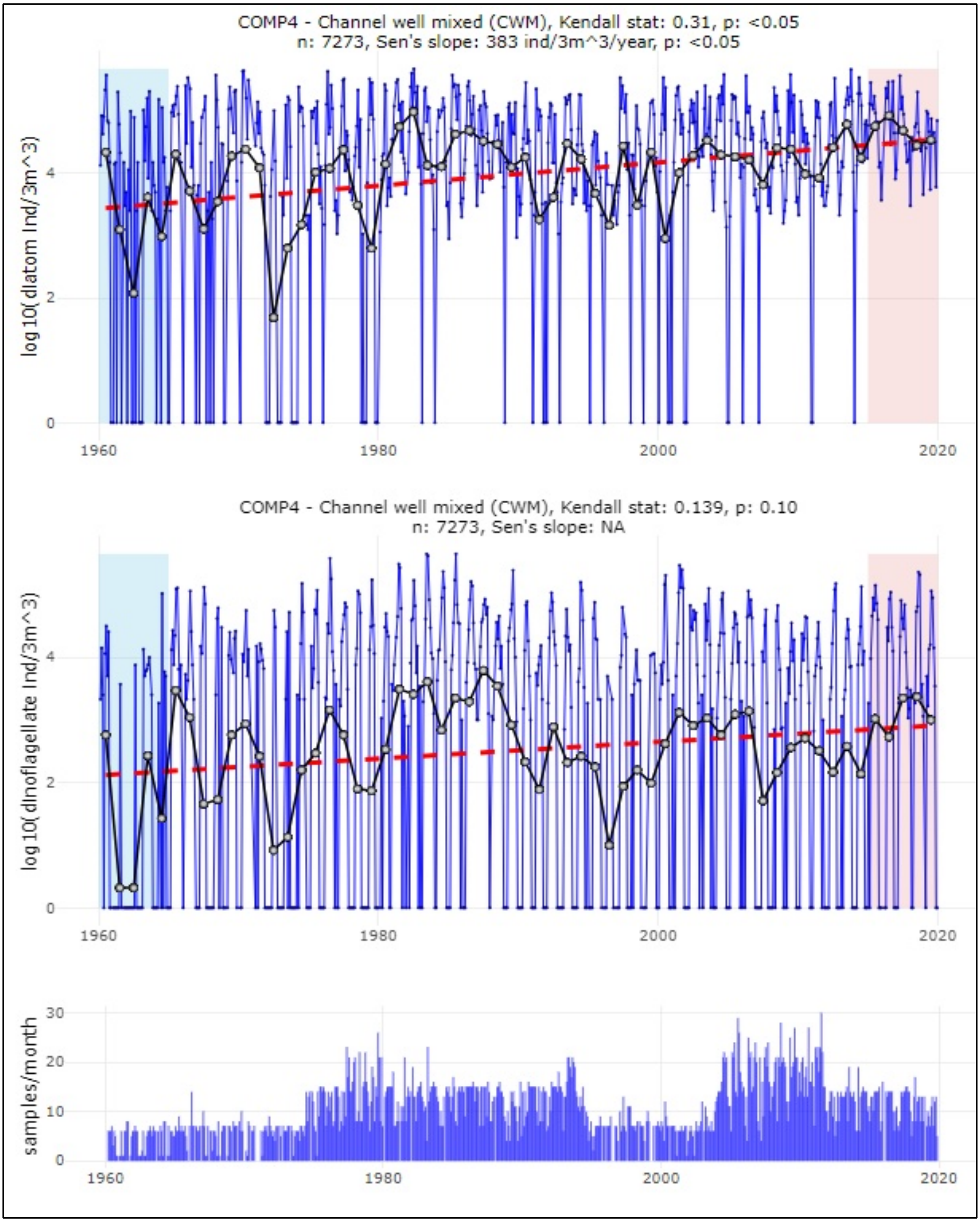
Once a dataset has been selected, the user must select a relevant lifeform pair from the **Select lifeform pair** drop-down menu. This list contains all the lifeform pairs that exist in the selected dataset.



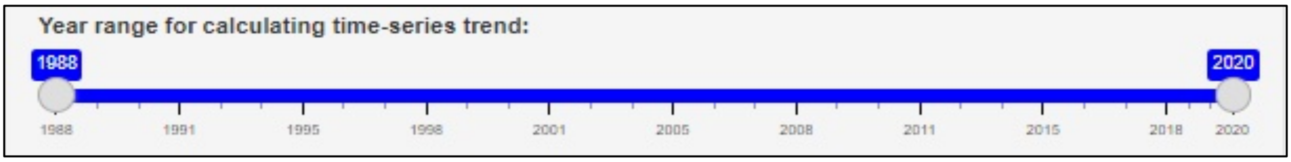
Once the lifeform pair is selected, the user can either select an assessment unit (spatial polygon or point, dependent on dataset type) from the **Select assessment unit** drop-down menu, or by clicking a polygon or point in either of the two map panels.



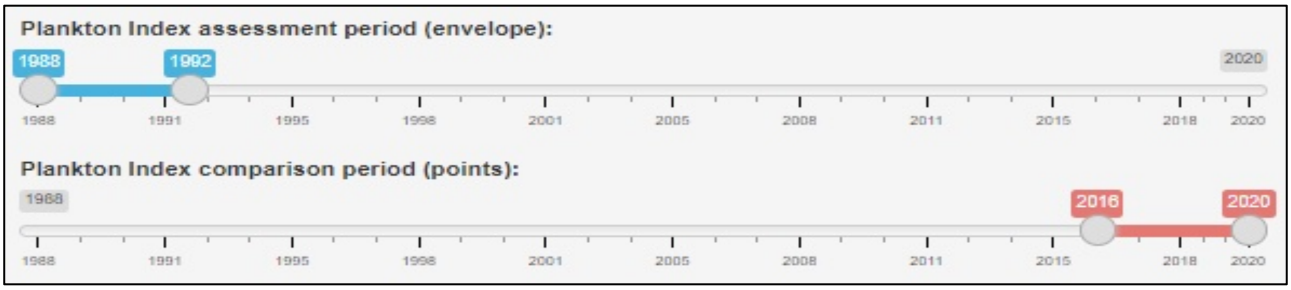
Once an assessment unit is selected, abundance time-series for the two lifeforms are rendered and Kendall statistics are calculated for the full time-series, indicating the direction of change over time, with negative numbers indicating decrease and positive indicating increase. This information is displayed above each time-series plot, along with Sen's slope estimate, describing the mean rate of change over time only for cases when the Kendall statistic is significant (i.e. p<0.05). The distribution of samples through time is also displayed as a separate plot below the two abundance time-series.



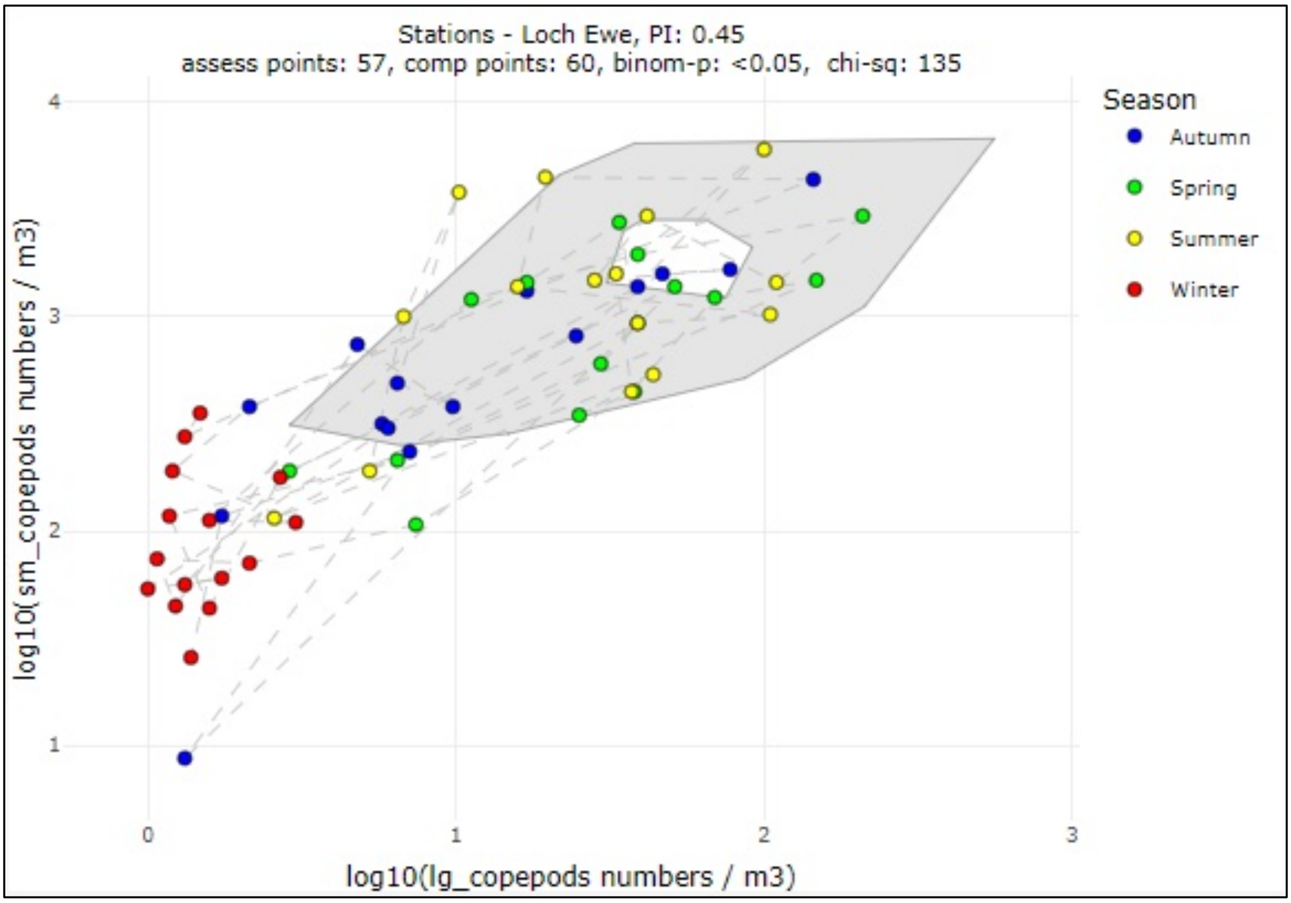
The length of the time-series assessed can be modified by adjusting the bounds of the slider labelled **Year range for calculating time-series trend**.



The shaded periods on each time-series indicate the **assessment period** used to calculate the plankton index (PI) envelope (blue shading) and the comparison period (pink shading). Both periods can be changed by adjusting the **Plankton Index assessment period (envelope)** and the **Plankton Index comparison period (points)** sliders.



Monthly abundance values from the assessment period (blue) are used to derive a 95% confidence polygon or "envelope" representing the mean annual cycle of covariance between the two lifeforms. Data from the comparison period (pink) are overlaid on top of this envelope to visualise how they differ from the assessment period.



Diagnostic information below the map panels indicates the confidence in the plankton index (PI) calculation. In some cases, the PI cannot be calculated due to insufficient data or overlap between the comparison and assessment periods. When this occurs, diagnostic messages inform why the indicator cannot be calculated. In other circumstances, the PI can be calculated, but only with low confidence due to issues revealed to the user in the diagnostic text.

The data are insufficient or unsuitable for calculating meaningful PI statistics.
The number of months in the assessment period does not fit the minimal criteria to generate the assessment envelope (i.e. n < 10).
The number of months in the assessment period does not fit the robust criteria (i.e. n < 30).
The number of years in the assessment period does not fit the robust criteria (i.e. n < 3).

Finally, COMP4 and WFD layers can be switched on or off using the layers tab in the top right of each map pane.

Appendix 4. Zooplankton statistical modelling

Sampling and data preparation

We have used a model-based approach to explore underlying trends in the zooplankton abundance data and how these relate to water quality parameters. The underlying data have been gathered since mid-2022 as part of the Environment Agency's monitoring programme. As these samples were gathered as part of the Environment Agency's routine water quality monitoring program, each zooplankton sample has a water quality sample associated with it.

The zooplankton data were returned from the laboratory with zooplankton quantities expressed as counts of individuals. As there was variability in the volume of water sampled, these quantities were standardised to zooplankton densities per m^3 by dividing counts of each taxon by the estimated volume of water sampled.

Water quality data were extracted from the Environment Agency's Water Information Management System (WIMS) data base. Zooplankton density data were matched to the WIMS sample data through cross-referencing with the associated PRN value, which provide a unique reference for each sampling event. The following water quality parameters were extracted from the WIMS data:

- Ammonium (NH_4) concentration (mg/l)
- Dissolved inorganic nitrogen (DIN) concentration (mg/l)
- Orthophosphate (PO_4) concentration (mg/l)
- Salinity (ppt)
- Chlorophyll *a* concentration (µg/l)
- Water depth (m)
- Water temperature (°C)

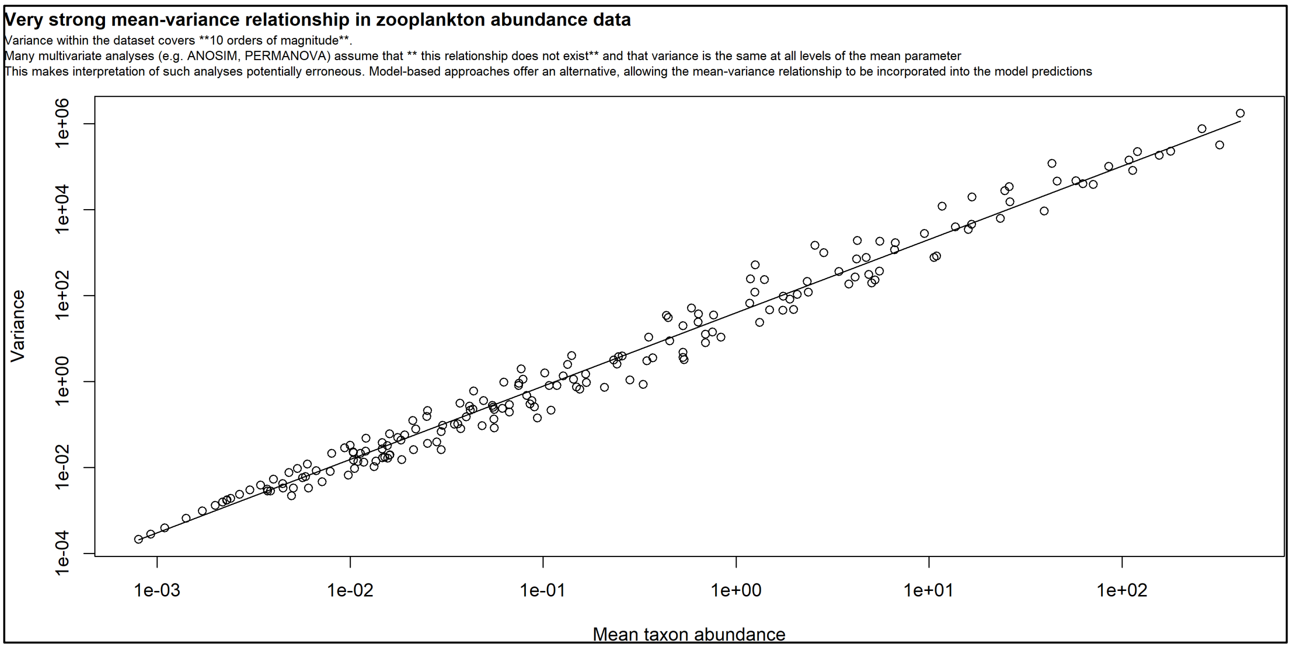
This resulted in a sample-by-taxon table and a sample-by-water quality table. These data were analysed using a generalised linear latent variable model (using the R package "gllvm", Niku et al., 2019).

Prior to analysis, water quality parameter values recorded at lower than the analytical equipment's limit of detection (LOD) were replaced with values of 0.5*LOD. For example, values recorded as "<0.1 mg/l" were recorded as 0.5*0.1 = 0.05 mg/l. Furthermore, as the gllvm software does not allow for missing values in the analyses, missing values in the WIMS data were replaced with the mean value for the given variable. Using the mean value minimises the effect that these 'missing' values have on the model calculations.

Generalised linear latent variable model approach

Traditionally, approaches such as analysis of similarities (ANOSIM) and permutational multivariate analyses of variance (PERMANOVA) have been utilised to describe the underlying structure of multivariate species assemblage data. These are largely *ad hoc* approaches, and have a number of issues associated with them, including low statistical power, and assumptions that are unlikely to be met in real world ecological data[1]. In the last ~10 years there has been an increasing move towards utilising model-based approaches for the assessment of such data.

One reason for this shift towards model-based approaches is that distance-based approaches do not account for the strong mean-variance relationship common to ecological data. That is, the models underlying these traditional approaches are based on the assumption of homogenous variability, regardless of mean taxon abundance. However, this is not the case, with variance levels clearly showing a strong positive relationship with mean taxon abundances:



A mixed effects model approach was undertaken, with water quality variables modelled as fixed effects. The geographic region of the water body sampled was modelled as a random effect. As the discrete zooplankton abundance count data were converted to continuous density data, the model adopted a Tweedie distribution. This distribution allows continuous values between zero and infinity.

Lifeforms data

Taxon names in and of themselves provide only a partial insight into ecosystem functioning and ecological processes. That is, the arbitrary labels we assign to species provide no tangible insight into the ecological roles that taxa play within an ecosystem. A complimentary approach to using taxon names is to assign taxa to ecologically-important traits or lifeforms.

- Taxa assigned to lifeforms
- Abundances of shared lifeforms summed within each sample
- Analyses repeated using lifeforms instead of taxon names

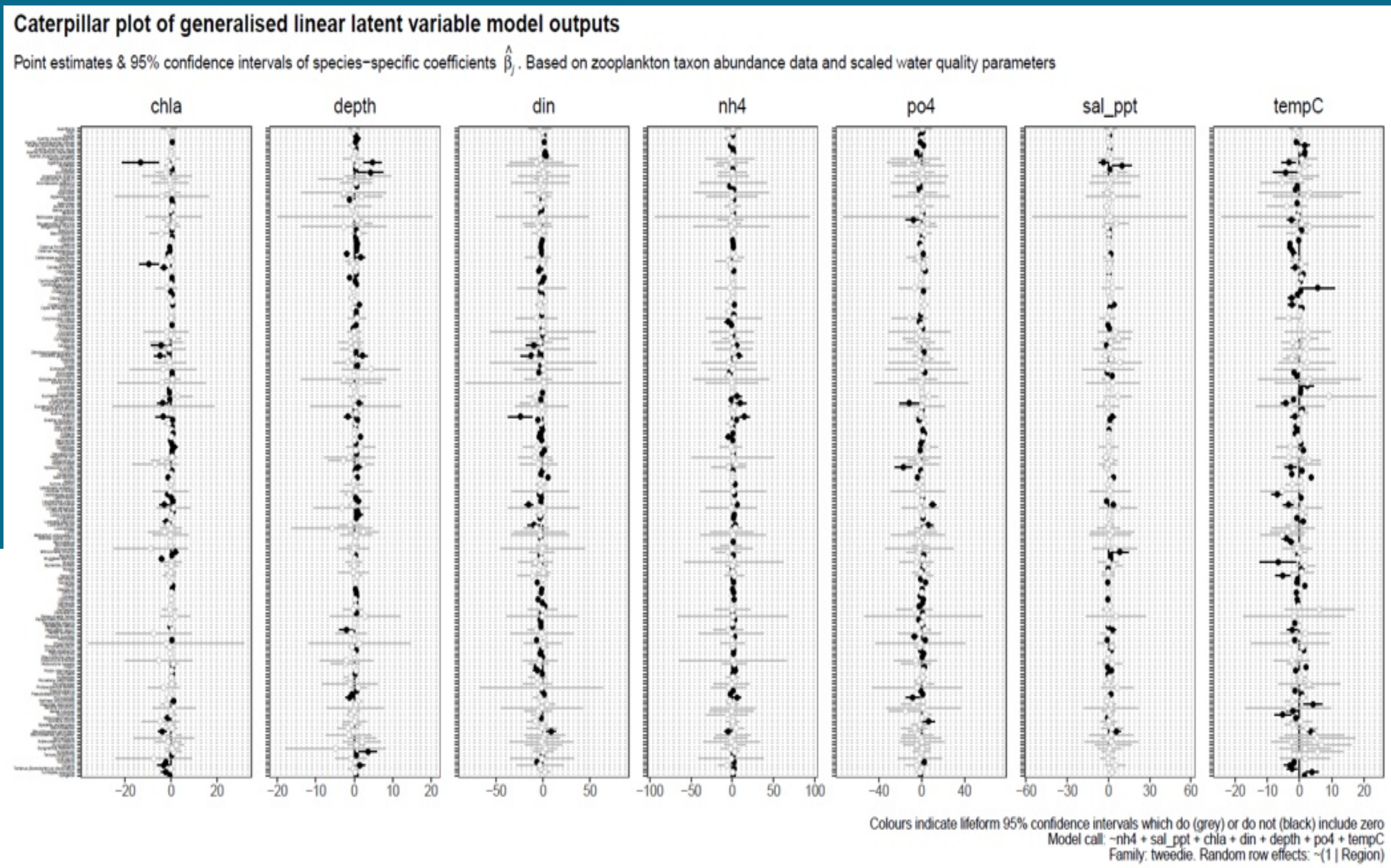
[1] See Warton, D. I., Wright, S. T., & Wang, Y. (2012). Distance-based multivariate analyses confound location and dispersion effects. Methods in Ecology and Evolution, 3: 89-101.

GLLVM outputs

To graphically display the gllvm model outputs, the coefficients of the first latent variable were extracted for each taxon or lifeform. This comprised a point estimate and a 95% confidence interval. These were then plotted, with taxon/lifeforms whose confidence intervals did not cross over zero highlighted. In the "caterpillar plots" below:

- confidence intervals which **do** include zero (i.e., the grey ones) implies a degree of uncertainty about whether the 'true' association of a given taxon/lifeform with the variable in question is positive or negative. We generally take this to mean there is no statistically clear evidence that the effect of the variable in question (e.g., chla, depth, DIN, etc.) is different from zero. As such, the variable might not have a clear impact on the outcome for that particular lifeform/species.

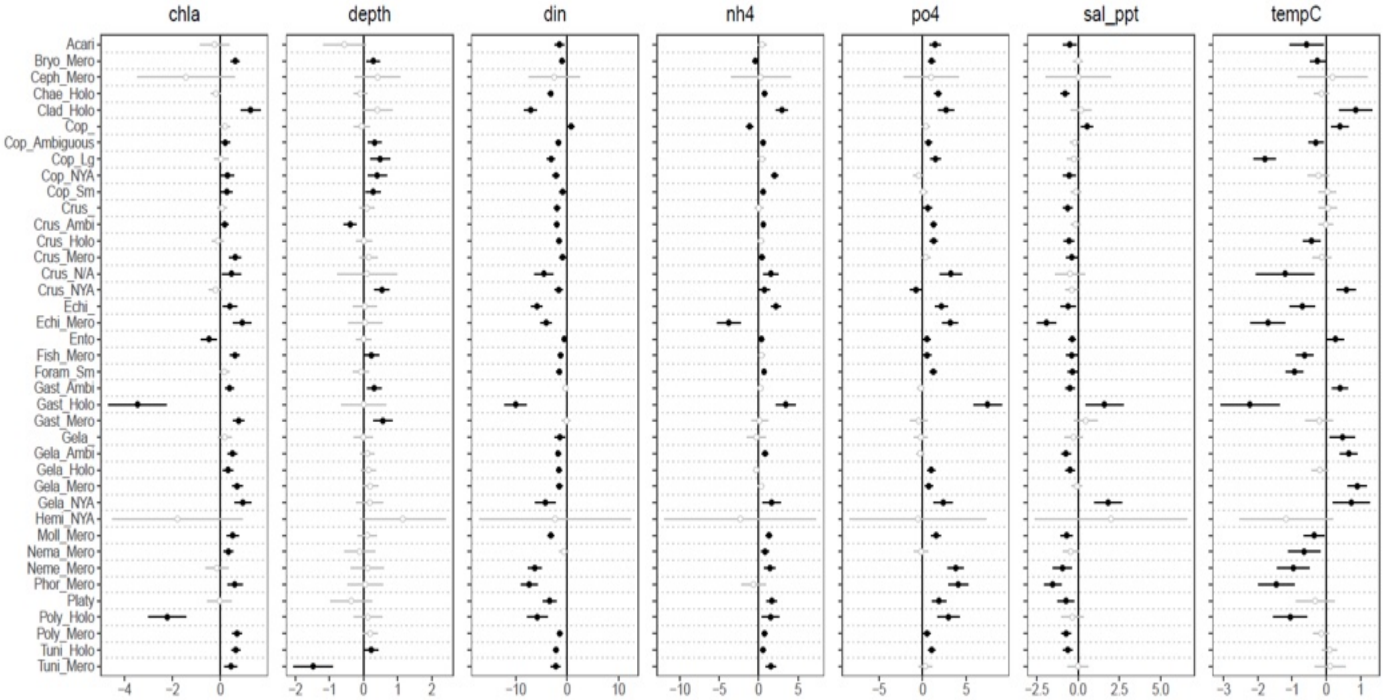
- confidence intervals which **do not** include zero (black) implies that we are confident of a statistically clear difference from zero in our parameter estimate. We can consider that we have 'statistically significant' evidence of association of the variable on our species/lifeform.



Caterpillar plot of generalised linear latent variable model outputs
Point estimates & 95% confidence intervals of species−specific coefficients $\hat{\beta}_j$. Based on zooplankton taxon abundance data and scaled water quality parameters
chla
depth
din
nh4
po4
sal_ppt
tempC
Colours indicate lifeform 95% confidence intervals which do (grey) or do not (black) include zero
Model call: ~nh4 + sal_ppt + chla + din + depth + po4 + tempC
Family: tweedie. Random row effects: ~(1 | Region)

Caterpillar plot of generalised linear latent variable model outputs

Point estimates & 95% confidence intervals of lifeform−specific coefficients $\hat{\beta}_j$. Based on zooplankton taxon lifeforms and scaled water quality parameters



Colours indicate lifeform 95% confidence intervals which do (grey) or do not (black) include zero
Lifeforms recorded in fewer than 5 samples removed from data prior to model estimations
Model call: ~nh4 + sal_ppt + chla + din + depth + po4 + tempC
Family: tweedie. Random row effects: ~(1 | Region)