

Supplementary Materials

Appendix A: Nutrient inputs overview

Overview of all rivers included in the exercise and their respective reduction percentages.

Supplementary Table 1: Overview of 1900 loads (estimated percentage of current day loads) as applied in the study. Reductions based on JMP-EUNOSAT results (in black) and local studies (in red). Where rivers were estimated to have higher loads or discharges around 1900 than currently the values applied were reset to 100%, i.e. the same as current day loads or discharges. The reduction percentages were applied to the loads in the observations-based OSPAR ICG-EMO riverine database.

River name	Country	Discharge %	HS TN%	HS TP%	River name	Country	discharge%	HS TN%	HS TP%
Meuse	NL	102	38	38	CARRADALE	GB	101	94	182
Rhine	NL	101	43	38	CARRON	GB	100	146	634
North Sea Canal	NL	106	30	27	CEFNI	GB	98	38	67
Lake IJssel East	NL	106	22	36	CHELMER	GB	118	52	65
Lake IJssel West	NL	104	21	36	CLEDDAU	GB	99	35	59
Schelde	NL	103	46	81	CLWYD	GB	101	41	66
GentTerneuzenCanal	NL	100	31	33	CLYDE	GB	99	43	45
Brede	DK	112	32	36	COBER	GB	105	55	54
Brons	DK	112	32	36	COLNE	GB	102	38	27
Elling	DK	111	28	36	CONDOR	GB	98	57	78
Gera	DK	116	30	36	CONON	GB	100	111	419
Grona	DK	116	29	36	CONWY	GB	98	56	60
Gudena	DK	111	26	36	COQUET	GB	101	51	95
Haslevgards	DK	116	24	36	Cormarty Firth	GB	102	61	94
Jordbro	DK	109	28	36	CREE	GB	101	73	160
Karup	DK	109	28	36	CUCKMERE	GB	98	40	36
Kastbjerg	DK	113	37	36	CURRYPOOL STREAM	GB	97	30	58
Konge	DK	117	29	36	DARENT	GB	98	59	80
Lindborg	DK	113	28	36	DART	GB	104	42	25
Liver	DK	116	34	36	DEBEN	GB	100	47	100
Omme	DK	110	30	36	DEE AT ABERDEEN	GB	101	66	152
Ribe	DK	116	30	36	DEE AT CHESTER	GB	99	49	57
Ry	DK	111	17	36	DEE AT TONGLAND	GB	101	66	122
Simested	DK	109	28	36	DENE MOUTH	GB	101	26	28
Skals	DK	109	28	36	DERWENT	GB	102	28	19
Skjern	DK	110	30	36	DEVERON	GB	103	45	68
Sneum	DK	116	31	36	DIGHTY WATER	GB	107	50	89
Stora	DK	108	32	36	DON	GB	103	50	74
Uggerby	DK	111	33	36	DONIFORD STREAM	GB	103	39	54
Varde	DK	114	32	36	Dornach Firth	GB	100	102	376
Vida	DK	117	30	36	DOUGLAS	GB	101	42	56
Voer	DK	116	30	36	DOVER	GB	104	59	60
Warnow	D	102	32	118	DWYFAWR	GB	99	97	277
Trave	D	105	29	96	DWYRYD	GB	99	97	277
Schwentine	D	105	31	98	DYFI	GB	99	65	216
Elbe	D	105	51	26	DYSYNNI	GB	100	84	331
Ems	D	106	26	17	EARN	GB	101	54	94
Weser	D	104	37	24	EAST SOLENT	GB	99	41	53
Eider	D	101	23	8	EBBW FAWR	GB	96	48	193
Miele	D	101	25	83	EDEN AT CARLISLE	GB	101	48	68
Pinnau	D	103	27	77	EDEN IN SCOTLAND	GB	107	50	89

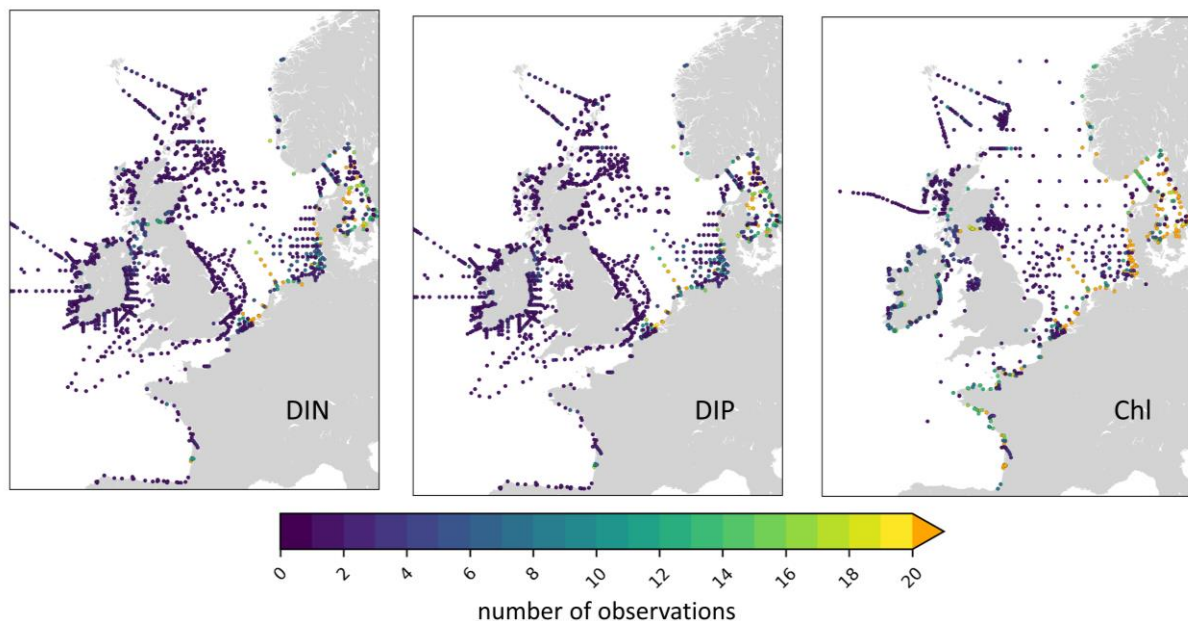
Tetenbuellspieker Kanal	D	101	23	73	EHEN	GB	100	16	10
Accumersiel	D	100	26	57	ELLEN	GB	100	58	117
Bensersiel	D	100	26	57	ELY	GB	97	23	15
Harlesiel	D	99	23	73	Enler	GB	100	29	21
Knock	D	101	23	68	ERCH	GB	101	42	71
Leybuchtziel	D	100	26	57	ERME	GB	101	45	50
Neuharlingersiel	D	99	23	73	ESK AT EDINBURGH	GB	101	26	25
Dangaster Siel	D	97	30	72	ESK AT WHITBY	GB	106	35	45
Eckwarder Siel	D	96	20	89	ESK INTO SOLWAY FIRTH	GB	101	62	132
Ems-Jade Kanal	D	97	30	72	EWE	GB	100	138	413
Fedderwardsiel	D	95	26	95	EXE	GB	102	41	51
Jade-Wapeler Siel	D	96	20	89	EYE WATER	GB	103	56	83
Maade Siel	D	97	30	72	FAL	GB	105	55	54
Schleuse Hooksiel	D	95	26	95	FINDHORN	GB	101	94	270
Schweiburger Siel	D	96	20	89	Finn	IRL	101	42	69
Vareler Siel	D	96	20	89	FIRTH OF FORTH	GB	99	58	103
Wanger Siel	D	100	25	45	FOWEY	GB	105	45	55
Arlau	D	110	26	56	Foyle	GB	101	42	69
Bongsieler Kanal	D	110	26	56	FROME	GB	100	39	52
Husumer Au	D	104	25	70	GIPPING	GB	114	49	82
Deichsiel Suederhafen	D	101	25	83	GIRVAN	GB	100	60	5
Gota alv	S	102	56	62	GLASLYN	GB	99	97	277
Lagan	S	101	48	57	GREAT EAU	GB	116	33	15
Nissan	S	100	48	45	GREAT STOUR	GB	106	49	44
Ronnean	S	101	35	57	GWYRFAI	GB	98	38	67
Viskan	S	101	47	62	GYPSEY RACE	GB	107	34	43
Atran	S	101	48	66	HALLADALE	GB	100	98	400
Deba	E	99	44	34	HASTINGS	GB	94	58	118
Oria	E	99	39	43	HAYLE	GB	107	49	49
Oiartzun	E	99	31	21	HELMSDALE	GB	100	96	377
Urola	E	99	44	34	HOLLAND BROOK	GB	112	40	43
Urumea	E	99	31	21	HUMBER	GB	110	34	33
Scheldt	BE	103	46	81	ILFRACOMBE	GB	103	39	54
SchipdonkCanal	BE	100	25	49	INVER	GB	100	148	435
BlankebergseVaart	BE	100	17	76	IRVINE	GB	99	20	7
LeopoldCanal	BE	100	25	49	ITCHEN	GB	99	51	86
IJzer	BE	107	23	61	KENT	GB	98	42	48
GentOostendeCanal	BE	100	17	76	Lagan	GB	100	42	104
Seine	FR	104	45	71	LEVEN	GB	98	53	89
Authie	FR	101	35	79	LEVEN IN SCOTLAND	GB	104	29	24
Canche	FR	109	44	63	LITTLE EACHAIG	GB	101	78	164
Somme	FR	106	27	86	LOCHY	GB	100	102	484
Loire	FR	100	50	92	LOSSIE	GB	101	51	48
Vilaine	FR	103	55	76	LOUGHOR	GB	97	39	77
Garonne	FR	107	70	74	LUD	GB	116	33	15
Charente	FR	107	62	88	LUNE	GB	98	57	78
Dordogne	FR	103	57	82	LYMN	GB	125	43	45
Eyre	FR	105	72	45	MAWDDACH	GB	99	40	42
Orne	FR	100	46	38	MEDWAY	GB	102	39	39
Seudre	FR	100	35	17	MEON	GB	99	41	53
Sevre-Niortaise	FR	107	45	91	MERSEY	GB	98	37	39
Shannon	IRL	100	29	53	Mourne	GB	101	42	69
Corrib	IRL	100	39	75	NAIRN	GB	102	30	13
Erne	IRL	102	38	62	Naver	GB	100	122	431
Ballysadare	IRL	102	35	66	NEATH	GB	99	30	51
Blackwater	IRL	102	35	55	NENE	GB	119	37	49
Moy	IRL	102	39	72	NESS	GB	101	107	335
Suir	IRL	102	34	57	NEVIS	GB	100	102	484
Barrow	IRL	102	34	57	Newry	GB	100	25	29
Nore	IRL	102	34	57	NITH	GB	100	56	109
Boyne	IRL	102	31	50	NORTH ESK	GB	103	62	100
Slaney	IRL	104	34	57	OGMORE	GB	99	30	51

Feale	IRL	101	42	63	OGWEN	GB	99	50	61
Garavogue	IRL	102	42	72	ORCHY	GB	100	87	331
Mulkear	IRL	100	29	53	OTTER	GB	102	41	51
Clarín	IRL	100	41	84	OUSE AT KINGS LYNN	GB	119	37	49
Bandon	IRL	102	33	45	OUSE AT NEWHAVEN	GB	98	40	36
Maigue	IRL	100	37	65	PARRETT	GB	97	30	58
Fergus	IRL	101	43	62	PENZANCE	GB	105	55	54
Deel	IRL	101	33	56	PIDDLE	GB	100	39	52
Liffey	IRL	100	36	74	PLYM	GB	102	49	56
Dunkellin	IRL	100	41	84	PORLOCK BAY	GB	103	39	54
Glyde	IRL	104	29	55	Quoile	GB	100	40	65
Dee	IRL	104	29	55	RED	GB	107	49	49
Avoca	IRL	103	34	54	RHEIDOL	GB	98	46	91
Lee	IRL	102	33	35	RHYMNEY	GB	99	46	74
Fane	IRL	102	34	62	RIBBLE	GB	99	38	32
Inagh	IRL	101	45	66	Roe	GB	100	35	48
Laune	IRL	101	40	45	ROTHER	GB	97	56	89
Leannan	IRL	102	44	59	SCARBOROUGH	GB	107	34	43
Maine	IRL	100	39	55	SEVERN	GB	103	42	54
Nanny	IRL	102	31	50	SHIEL	GB	100	114	440
Owenavorrhagh	IRL	103	34	54	SHIN	GB	100	102	376
Owenboliskey	IRL	102	40	66	SOUTH ESK	GB	103	62	82
Owenboy	IRL	103	32	32	SPEY	GB	101	82	237
Owenea	IRL	101	44	61	STOUR AT BOURNEMOUTH	GB	102	41	50
Tolka	IRL	102	14	9	STOUR AT HARWICH	GB	122	53	50
Swilly	IRL	102	44	59	TAF	GB	99	37	58
Glomma	NO	101	44	50	TAFF	GB	99	46	74
Numedal	NO	100	53	64	TAMAR	GB	105	45	55
Skien	NO	100	47	76	TAW	GB	102	39	55
Drammen	NO	103	25	53	TAWE	GB	97	44	60
Otra	NO	100	48	91	TAY	GB	101	63	143
Orre	NO	102	33	46	TEES	GB	101	26	28
bjerkreim	NO	101	33	40	TEIFI	GB	98	37	69
elkelandsosen	NO	100	47	32	TEIGN	GB	104	47	60
kvina	NO	100	37	80	TEST	GB	102	36	49
lygna	NO	100	37	58	THAMES	GB	100	35	38
mandal	NO	101	44	59	THAW	GB	97	37	57
nidelva trondheim	NO	101	63	64	THEDDLETHORPE	GB	122	34	33
nidelva	NO	100	53	91	THURSO	GB	101	91	244
orkla	NO	101	63	64	TIDDY	GB	105	45	55
sira	NO	100	21	8	TORRIDGE	GB	102	39	55
suldal	NO	100	39	62	TWEED	GB	103	56	83
tovdal	NO	100	32	66	TYNE	GB	100	47	43
gaula	NO	101	63	64	TYNE IN SCOTLAND	GB	110	44	53
ADUR	GB	99	50	80	TYWI	GB	98	40	76
AFON GOCH	GB	98	38	67	UGIE	GB	101	40	65
ALDE	GB	100	47	100	URR WATER	GB	101	47	83
ALMOND	GB	100	49	55	USK	GB	99	46	74
ALT	GB	99	42	73	WALLERS HAVEN	GB	95	47	61
ANNAN	GB	101	52	86	WALLINGTON	GB	99	41	53
ARUN	GB	99	57	60	WANSBECK	GB	106	32	29
AVON AT BANTHAM	GB	104	42	25	WATER OF LEITH	GB	101	26	25
AVON AT BOURNEMOUTH	GB	102	41	50	WATER OF LUCE	GB	99	97	316
AVON AT BRISTOL	GB	100	41	71	WAVENEY	GB	118	36	38
AXE	GB	103	38	36	WEAR	GB	101	44	64
AYR	GB	100	51	100	WEAVER	GB	103	36	48
BABINGLEY	GB	119	37	49	WELLAND	GB	119	41	38
Bann	GB	100	35	48	WEST BAY	GB	103	39	56
BEAULY	GB	101	107	335	WEY	GB	103	39	56
BLYTH	GB	106	32	29	WICK	GB	101	71	134
BLYTH AT WALBERSWICK	GB	114	41	52	WITHAM	GB	124	44	53

BRIDLINGTON	GB	107	34	43	WYE	GB	102	48	61
BRORA	GB	101	76	172	WYRE	GB	98	57	78
BURE	GB	118	36	38	YAR	GB	101	41	48
Bush	GB	100	35	48	YARE	GB	118	36	38
CADOXTEN	GB	97	37	57	YSTWYTH	GB	99	30	54
CAMEL	GB	107	54	51	YTHAN	GB	105	37	72
CARNON	GB	105	55	54					

Appendix B: Observations overview

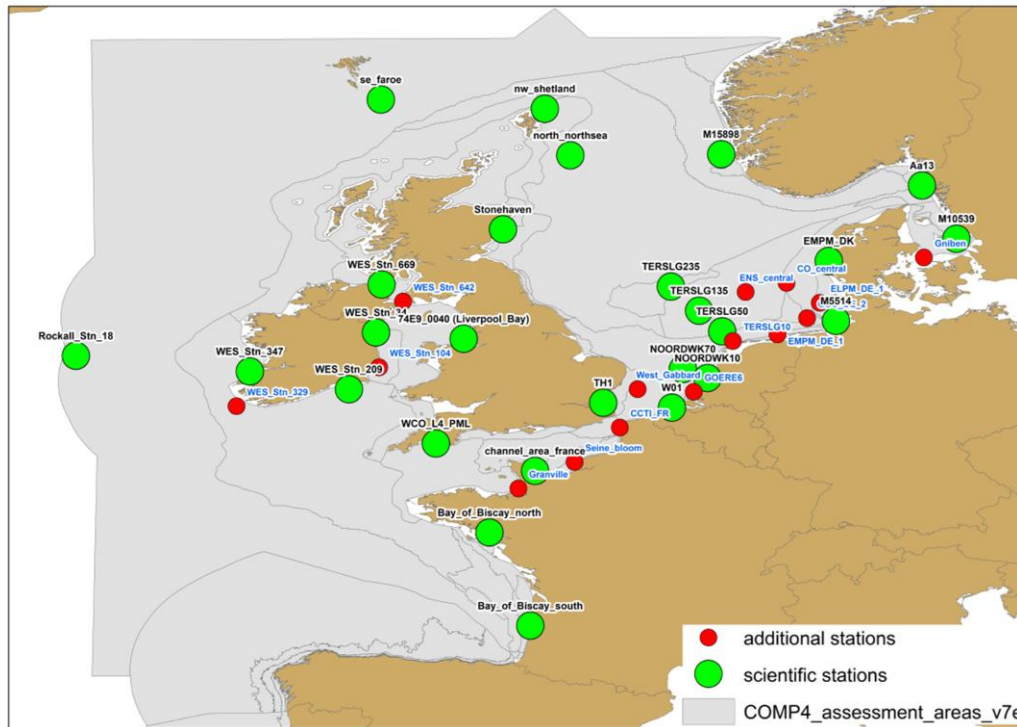
The weighting method applied in this study is based on observations. As such, the quality of the observations is of prime importance, and is considered here. First we display the spatial distribution of the observations used in the weighting method (Supplementary Figure 1), which were extracted from the ICES data base (ICES, 2022, <https://data.ices.dk/view-map>) using the COMPEAT tool. The weighting method compares the averaged simulated results with the averaged observations (spatially averaged per area, temporally averaged over 2009-2014), thus the spatial and temporal distribution of observations within the areas is important. For DIN and DIP the relative variance (i.e. standard deviation/mean) is generally higher for inshore assessment areas than offshore ones. The assessment areas have been chosen to have similar oceanographic properties, but inshore areas can have steep gradients in DIN/DIP. There may also be a bias in some areas where the sampling locations do not reflect the mean concentration over the entire area, resulting in a mismatched comparison to simulated results which do cover the entire area and period. Supplementary Figure 1 shows the spatial distribution of in situ observations used for DIN, DIP and Chl. As expected, most of the effort is focused on- the coastal zones, with a sparsity of *in-situ* observations for some offshore areas. This applies particularly to the Northern North Sea, Norwegian trench, much of the Atlantic and Bay of Biscay areas.



Supplementary Figure 1: Spatial distribution of the in-situ observations in COMPEAT for DIN, DIP and Chl in the period 2009-2014, as used in this study in the weighting method (Almroth & Skogen, 2010).

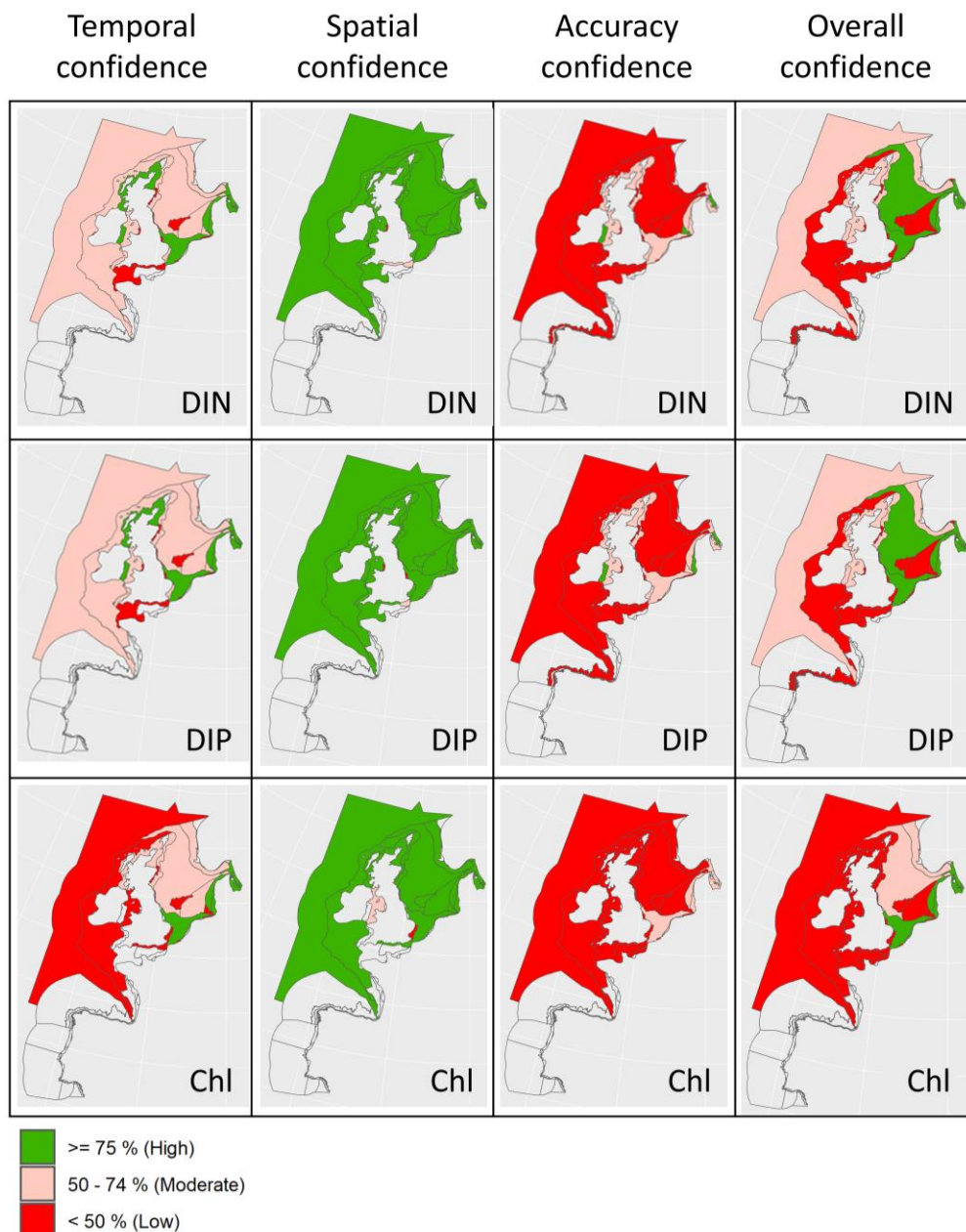
Secondly, we show the additional stations where observations were gathered for the model validation exercise (Supplementary Figure 2, see Figure 2 main article). For this purpose

stations were selected that had continuous measurements throughout the simulated time period. Note again the coastal bias present in the stations' spatial distribution.



Supplementary Figure 2: stations used in the common validation exercise based on long-term observational time series.

The COMPEAT tool provides 3 methods of assessing observational confidence: temporal distribution, spatial distribution and accuracy (function of how close the mean results are to a threshold) for each parameter. These are then combined in an overall confidence level (Supplementary Figure 3). For DIN and DIP there is poor temporal confidence for all of the offshore regions. Spatial confidence is generally better, but overall confidence is only high for regions in the North Sea. In general there is lower confidence in the *in-situ* chlorophyll observations than in those of DIN or DIP.

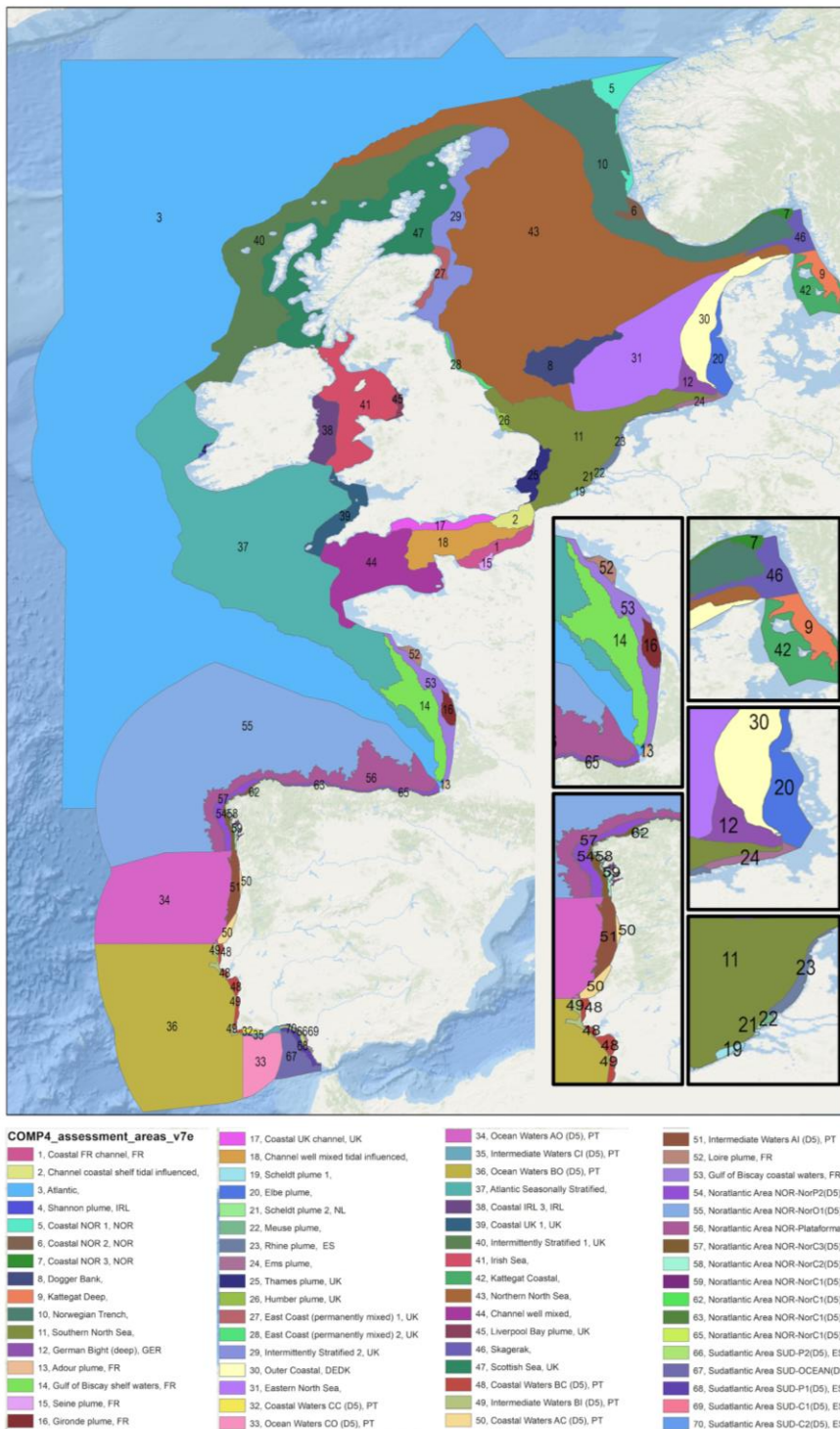


Supplementary Figure 3: From left to right: temporal distribution, spatial distribution, accuracy and overall confidence assessments of the applied *in situ* observations as derived by COMPEAT, for DIN, DIP and Chl.

Because of the poor confidence in *in-situ* chlorophyll measurements, satellite derived values were incorporated into the COMPEAT tool. This result of the combined satellite – *in-situ* product gives high confidence for temporal, spatial, accuracy and overall. It is this combined value which has been used in the weighting. The consequence of using this combined product is that the relative variance drops to around 13 % for chlorophyll, however, there is some uncertainty and potential bias in near shore chlorophyll satellite measurements.

Appendix C: COMP4 areas

Supplementary Figure 4 shows the areas as defined by OSPAR for use in the 4th application of the Common procedure (COMP4, Enserink et al, 2019), based on the work of van Leeuwen et al (2015). Note that not all areas are covered by the participating models: only 2 models include the major part of the Bay of Biscay and none cover the Portuguese coastline.



Supplementary Figure 4: Overview of the COMP4 assessment areas (version 7e, numbering by ICG-EMO).

Appendix D: Model overview

ECOHAM (University of Hamburg and HZG, Germany)

ECOHAM (ECOLOGICAL Model-HAMBURG) is defined on 31 z-levels and a $1/5^{\circ} \times 1/3^{\circ}$ km grid, that covers almost the entire Northwest Continental Shelf. The biogeochemical model consists of two phytoplankton (diatoms and flagellates), two zooplankton groups (micro- and mesozooplankton) and a bacteria group with fixed stoichiometry, O₂, and C, N, P bound to two detritus size classes, dissolved organic material, dissolved inorganic material and 2D (plate) sediment pools. At the open ocean boundaries, biogeochemical model variables are nudged to the values extracted from the World Ocean Atlas for all scenarios. Transport of modelled biogeochemical variables are calculated based on the diffusion and convection fields estimated by the hydrodynamical model HAMSOM (HAMBURG Shelf Ocean Model). Details about the model setup and biogeochemical model can be found in Große et al., 2017 and references therein.

GPM (University of Oldenburg, Germany)

GPM (Generalized Plankton Model) is defined on 20 σ -layers and a 1.5-4.5 km curvilinear grid, covering only the southern North Sea. In the present implementation, the ‘geochemistry’ portion (O₂ and C, N, P bound to dissolved inorganic and organic material, detritus and sediment pools) of the model is as in ECOHAM, but for the description of plankton growth and interactions, the variable chlorophyll content and C:N:P ratios of phytoplankton were taken into account. At the open ocean boundaries, geochemical variables are clamped to the ECOHAM results obtained for respective scenarios, whereas for plankton variables, zero-gradient conditions were assumed. The biogeochemical model is on-line coupled to the Generalized Estuarine Transport Model (GETM) as the hydrodynamical driver. A detailed description of the model setup and the biogeochemical model can be found in Kerimoglu et al. 2020 and references therein.

MIRO&CO (RBINS, Belgium)

MIRO&CO results from the coupling of the 3D hydrodynamic COHERENS model (Luyten, 2011) with the biogeochemical MIRO model (Lancelot et al., 2005). COHERENS is a three-dimensional numerical model, designed for application in coastal and shelf seas, estuaries, lakes, reservoirs. MIRO is a biogeochemical model that has been designed for *Phaeocystis*-dominated ecosystems. It describes the dynamics of phytoplankton (three functional groups), zooplankton (two functional groups), heterotrophic bacteria, organic matter degradation (dissolved and particulate) and nutrient cycles (N, P, Si) in the water column and the sediment. The current setup has been obtained by coupling MIRO with COHERENS v2 (MIRO&CO v2): details and validation are shown in Dulière et al. (2019).

GETM-FABM-ERSEM (JRC Ispra, European Union)

The JRC-NWES is largely described by Friedland et al. (2020), including a variety of validation results. For the present study, the model setup was adapted to the river inputs provided within ICG-EMO and extended with a more sophisticated sediment model, inclusion of atmospheric nitrogen deposition and an improved attenuation calculation method.

ECO-MARS3D (IFREMER, France)

The ECO-MARS3D ecological model is based on the MARS3D hydrodynamical code (Lazure & Dumas, 2008), a three-dimensional model based on Navier-Stokes equations under the classic Boussinesq and hydrostatic assumptions within a sigma framework. The originality of this model is on the coupling between barotropic and baroclinic modes especially designed for the alternate direction implicit method (ADI). The time-step is adaptive and the model is forced by a barotropic sea-level oscillation (at the oceanic boundaries) and by atmospheric conditions (throughout the domain). It provides realistic descriptions of coastal hydrodynamics for research and operational interests. The fully coupled biogeochemical module ECO-MARS3D is a NPZD model type (Nutrient–Phytoplankton–Zooplankton–Detritus), that aims to simulate the fluxes of limiting elements such as nitrogen (N), phosphorus (P) and silicon (Si). The phytoplankton compartment is described by the following variables ‘diatoms’, ‘dinoflagellates’, ‘nanopicoplankton’ and the haptophyte ‘*Phaeocystis globosa*’. The grazers are split into two groups ‘micro-zooplankton’ and ‘meso-zooplankton’ and the detritus coming from phytoplankton, zooplankton senescence and excretion are mineralized and contribute to the nutrient renewal. Moreover, variables ‘dissolved oxygen’ as well as ‘oxygen saturation’ are calculated taking into account air-water exchanges, primary production and respiration processes in the water column. The current application to the French Atlantic shelf is based on a regular grid with 4×4 km meshes and 30 sigma layers, which covers the Bay of Biscay, the English Channel and the southern part of the North Sea, up to the Rhine estuary. Detailed description and validation is provided in Ménesguen et al. (2018, 2019).

DFLOW-FM (Deltares, the Netherlands)

The model used by Deltares for the model comparison is a combination of the Generic Ecological Model (GEM) for the water quality and ecological processes (Blauw et al., 2009) and a newly developed hydrodynamic model for the greater North Sea. The model uses the Delft Flexible Mesh (DFM) simulation software both for the hydrodynamics and water quality and ecological processes. The hydrodynamical part of the model is originally developed for flood forecasting and transport simulation purposes (Zijl et al., 2021). In combination with the GEM model it is used to study the effects of several anthropogenic impacts to the North Sea, such as eutrophication, climate change, aquaculture and wind farms. The GEM model simulates the nutrient cycles of nitrogen, phosphorus and silicate and the dynamics of phytoplankton and oxygen. Additionally, grazing by benthic filter feeders is included based on Dynamic Energy Budget (DEB) modelling (Troost et al., 2010; 2018). Four groups of phytoplankton are modelled (diatoms, flagellates, dinoflagellates and *Phaeocystis*) and 2 groups of benthic filter feeders (*Mytilus edulis* and *Ensis*).

NEMO-SCOBI (SMHI, Sweden)

The model used by the Swedish Meteorological and Hydrological Institute (SMHI) is based on the Nucleus for European Modelling of the Ocean (NEMO) framework (Madec, 2010), version 3.6 but specifically configured for the Baltic and North Seas (NEMO-Nordic; Hordoir et al., 2019). It has 56 vertical levels with a resolution of 3 m close to the surface and decreasing up to 22 m at the bottom of the deepest part of the domain (Norwegian trench). The horizontal resolution is of approximately 2 nautical miles (~3700 m). NEMO-Nordic has two open boundaries located in the English Channel between Brittany and Cornwall and between Scotland and Norway (Hordoir et al., 2019). The biogeochemistry is simulated by the Swedish Coastal and Ocean Biogeochemical module (SCOBI; e.g., Eilola et al., 2009). It includes the nutrient cycles of nitrogen, phosphorus and silicate in both the water column and sediments, simulates the oxygen cycle and accounts for 3 phytoplankton species in the water column: diatoms, flagellates and cyanobacteria (Almroth-Rosell et al., 2011).

All 3 runs with NEMO-SCOBI were initiated in year 1961 in order to ensure a close to balance Baltic Sea at years 1992-2014. Boundary profiles for nitrate, phosphate and oxygen used to force the model are similar to those from CMEMS. Atmospheric NO_x and NH_x input from EMEP MSC-W 2018 data (with agreed reduction for the Historical Scenarios) was used throughout the run. The river nutrient forcing provided by the ICG-EMO group for the historic scenario was adjusted to the model domain and combined to data of E-Hype version v.5.6.2. This former version of e-hype had to be used, as it showed important improvements in the salinity of the Baltic Sea when used as a forcing for NEMO-Nordic. Because this model domain includes the Baltic Sea and ICG-EMO river data did not include a reduction of P and N for the Historical Scenarios in this region, for these scenarios we used a constant factor to reduce each nutrient load in the Baltic Sea based on the reduction factors applied to the North Sea.

GETM-ERSEM-BFM (CEFAS, UK)

GETM-ERSEM-BFM, GETM (General Estuarine Transport Model) is a public domain, three-dimensional Finite Difference hydrodynamical model (www.getm.eu). It solves the 3D partial differential equations for conservation of mass, momentum, salt and heat, and was designed to handle drying and flooding (e.g. tidal flats). The ERSEM-BFM (European Regional Seas Ecosystem Model - Biogeochemical Flux Model) version is a development of the model ERSEM III (see Baretta et al., 1995; Ruardij and van Raaphorst, 1995; Vichi et al., 2007; van der Molen et al., 2013; www.nioz.nl/en/about/cos/ecosystem-modelling), and describes the dynamics of the biogeochemical fluxes within the pelagic and benthic environment. The ERSEM-BFM model simulates the cycles of carbon, nitrogen, phosphorus, silicate and oxygen and allows for variable internal nutrient ratios inside organisms, based on external availability and physiological status. The model applies a functional group approach and contains 6 phytoplankton groups, 4 zooplankton groups and 5 benthic groups, the latter comprising 4 macrofauna and 1 meiofauna group. Pelagic and benthic aerobic and anaerobic bacteria are also included. SPM concentrations are calculated as proportional to the local wave-induced bed-shear stress, varying linearly with depth, and with an exponential relaxation mechanism that represents delayed settling. The ERSEM-BFM model includes a 3-layer benthic module comprising 53 state variables. TEP (Transparent Exopolymeric Particles) production by diatoms is included, allowing for macro-aggregate formation and

rapid sinking out of the spring bloom. The ERSEM-BFM model also has enhanced pelagic-benthic coupling compared other ERSEM-III based models.

The setup includes a spherical grid covering the area 46.4°N-63°N, 17.25°W-13°E with a resolution of 0.08° in longitude and 0.05° in latitude (approximately 5.5 km), and 25 non-equidistant layers in the vertical. The model bathymetry was based on the NOOS bathymetry (www.noos.cc/index.php?id=173). The model was forced with tidal constituents derived from TOPEX-POSEIDON satellite altimetry and atmospheric forcing from ECMWF ERA-Interim.

Appendix E: Model comparison

Here a more direct comparison of the different model features is presented.

Supplementary Table 2: Direct comparison of model capabilities of all contributing marine ecosystem models

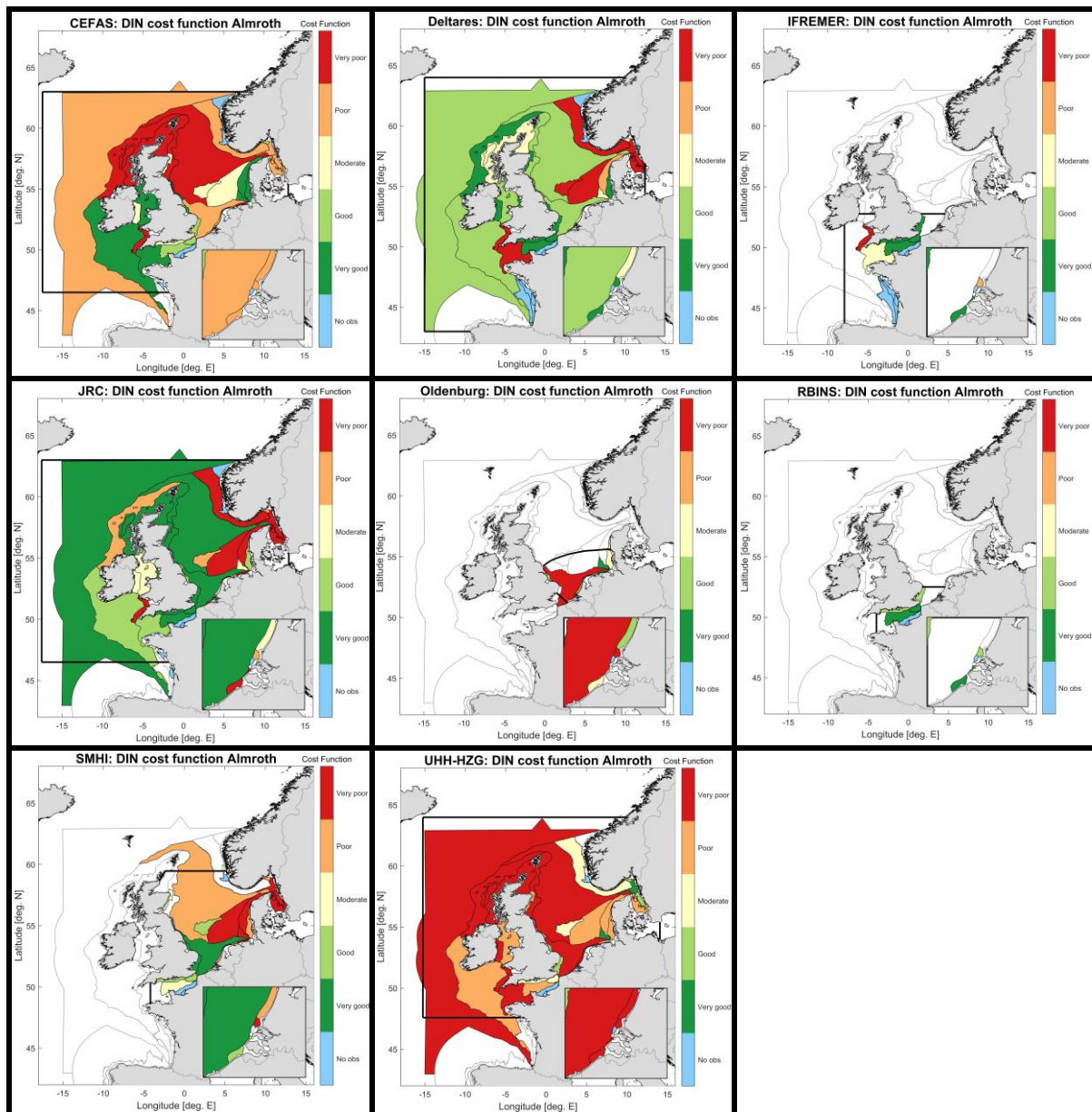
Model name	MIRO&CO-3D (RBINS, Belgium)	ECO-MARS3D (IFREMER, France)	EOHAM4 (UHH-HZG, Germany)	GPM (Oldenburg, Germany)	Def3D-GEM (Deltares, the Netherlands)	JRC-ERSEM (JRC Ispra, EU)	GETM-ERSEM-BFM (Cefas - United Kingdom)	NEMO -SCOBI (SMHI, Sweden)
General simulation characteristics								
Hydrodynamic model	COHERENS	MARS-3D	HAMSOM	GETM	DFLOW-FM	GETM	GETM	NEMO
Biogeochemical model	MIRO&CO	ECO-MARS-3D	EOHAM4	GPM	GEM	PML-ERSEM	ERSEM-BFM	SCOBI
Used model domain area	English Channel and Southern Bight of the North Sea	Southern North Sea, Channel, Celtic Sea, Bay of Biscay	European Shelf down to the top of the Bay of Biscay	Selected area of Southern North Sea	European Shelf including the entire Bay of Biscay	European Shelf halfway down to the Bay of Biscay	European Shelf halfway down to the Bay of Biscay	North Sea, Channel, Baltic Sea
Spatial Resolution Δh (km)	5.89 km (lon) x 4.63 km (lat)	4 km x 4km	20km x 20 km	1.5 – 4.5 curvilinear grid	1-8 km: 1 km in waters < 100 m deep, 8 km in water > 400 m deep and 4 km in between	4.04-6.13 km (x-direction) and 5.56 km (y)	5.5 x 5.5 km	~3.7 km
Vertical resolution	5 sigma layers	30 sigma layers	30 z layer	20 layers	20 layers	25 layers, dynamically adapting to density gradients	25 sigma layers	56 layers, with smaller layer width near the surface
Longitude (degree)	[-4.0,5.0]	-7.922° – 5.104°E	15.°W – 14.°E	0.15.°W – 9.15°E	-15 to + 14	17.5° W-13.1° E	17.25°W-13°E	4.15278° W to 30.1802° E

Model name	MIRO&CO-3D	ECO-MARS3D	ECOHAM4	GPM	Deft3D-GEM	JRC-ERSEM	GETM-ERSEM-BFM	NEMO -SCOBI
Latitude (degree)	[48.5,52.5]	52.769°N – 43.267°N	47.85°N – 64.°N	51.35°N – 55.6°N	44 to 62	46.4° N-63° N	46.4°N-63°N,	48.4917– 65.8914° N
Temporal resolution Δt (sec)	Hydrodynamics: 60s, MIRO: 15 min.	200-240	60 s	2D: 5 s 3D: 360 s	Hydrodynamics 2 min (max), water quality: 10 min.	12.87	15 seconds for hydrodynamics, 450 seconds for biology	Hydrodynamics: 3-90 s Biogeochemistry: 60 s
Temporal range	2000-2014	2006-2014	2009-2014	2009-2014	2009-2014	2006-2014	2009 – 2014	1992-2014
Spin up time	5 years	2006-2008	2006-2008	2006-2008	2006-2008	1 year	20 years	1961-1991
Meteo data	RMIB 6 hourly data based on UKMO atmospheric model.	Météo France ALADIN model	ERA 5	COSMO-CLM	ERA5	ERA5	ECMWF	UERRA
Inclusion of tides	Yes (15 harmonics)	Yes, through FES2004 tidal atlas (Lyard et al., 2006)	Yes, through sea surface elevation input from large-scale model	Yes, using hourly surface elevations estimated by TRIM-NP-2D : doi:10.1594/WDC C/coastDat-2_TRIM-NP-2d	Yes, through sea surface elevation input from CMEMS model	Yes. Using http://volkov.oce.orst.edu/tides/AO.html along the 2d boundaries	Yes, through sea surface elevation on the open boundaries from a larger scale model	Yes
Temperature & Salinity diagnostic or prognostic	Prognostic. SST from weekly sea surface gridded temperature (BSH)	prognostic	T: Prognostic S: diagnostic	T: prognostic S: prognostic	Prognostic i.e. simulated in the model	Prognostic	Prognostic	Prognostic

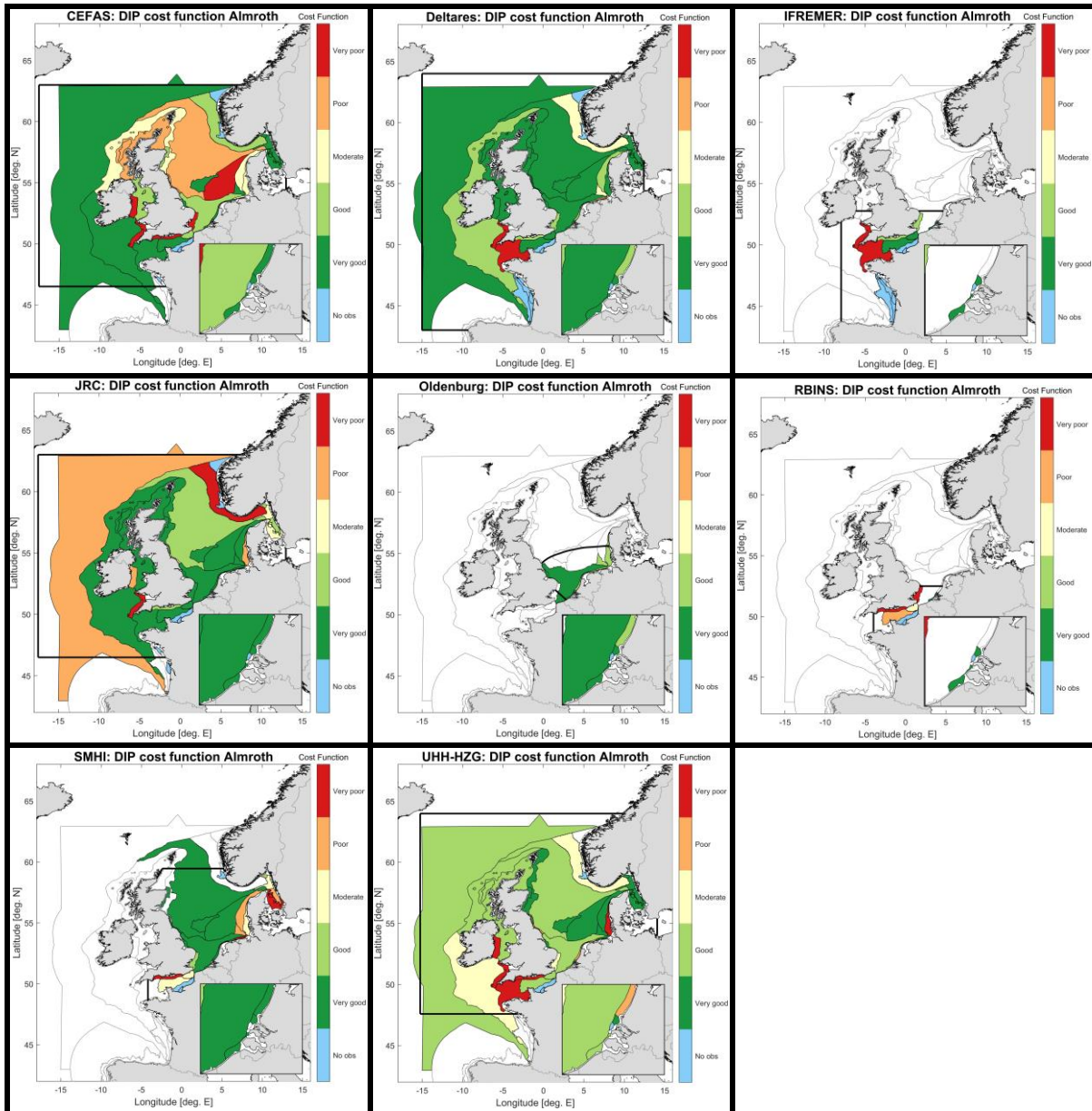
Model name	MIRO&CO-3D	ECO-MARS3D	ECOHAM4	GPM	Deft3D-GEM	JRC-ERSEM	GETM-ERSEM-BFM	NEMO -SCOBI
No. of benthic state variables	10	3	5	4	10 (4 detritus + 6 biota)	36	53	4 (PON, POP, Opal, inorganic phosphorous)
Benthic Nutrients (bulk or explicit)	Bulk	Bulk	Bulk	Bulk	Bulk (only detritus)	Explicit	Explicit	Explicit
Types of Zoobenthos	None	None	None	None	Mussels and Ensis	3 groups: benthic deposit feeder, benthic suspension feeder, benthic meiofauna	5 groups: meiobenthos, filter feeders, infaunal predators, deposit feeders, megabenthos	None
Types of benthic bacteria	None	None	None	None	None	2 groups: aerobic and anaerobic benthic bacteria	2 groups: aerobic and anaerobic benthic bacteria	None
Benthic DOM	No	No	No	No	No	Yes	Yes	No
Benthic POM	Refractory and non-refractory POC, PON, POP	Benthic organic matter in C, N, P, Si	No	No	One fraction of detritus for carbon, nitrogen, phosphorus and silicate	Benthic particulate organic matter , benthic refractory matter , benthic buried matter	Benthic POC, PON, POP	Yes, as particulate organic nitrogen (PON) and phosphorous (POP)
Participant to add further characteristics if required							TEP production by nutrient stressed diatoms included.	

Appendix F: Cost functions

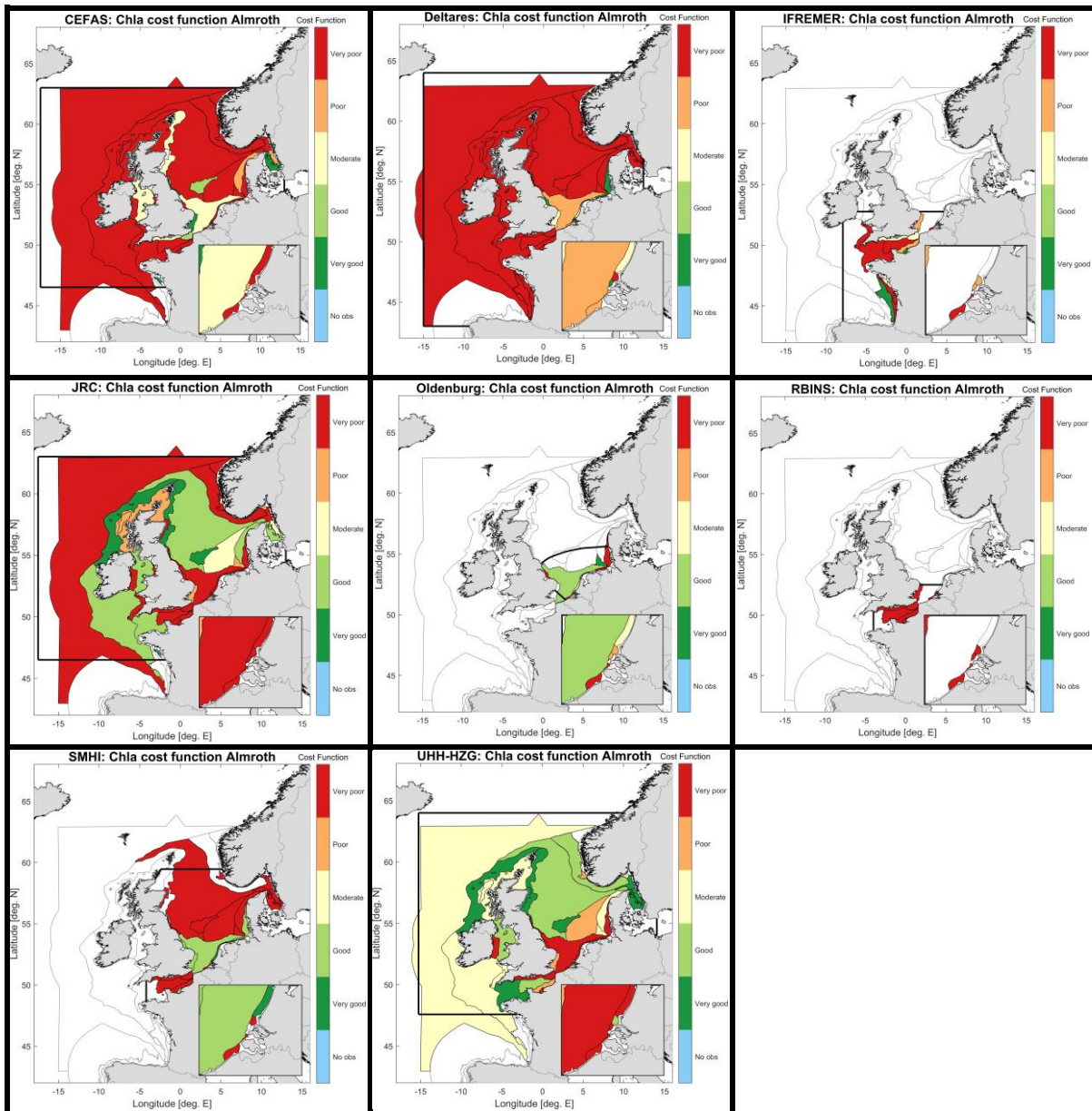
Here we show the individual results for the different models in terms of the cost function (eq. 1) from Almroth & Skogen (2010). The function scores have been replaced by general goodness-of-fit statements, as follows: 0-1 (very good), 1-2 (good), 2-3 (moderate), 3-4 (poor) 4-5 (very poor), with dark red areas representing cost function values > 5 . Note that a bad cost function score may be due to limited observations being available. The observations used in the cost function were extracted from the COMPEAT tool, and include both *in situ* and satellite observations for Chl.



Supplementary Figure 5: DIN cost function results from the Almroth & Skogen (2010) weighting method for all models. The thick lines indicate the domain boundaries for each participating model. Areas without sufficient model coverage ($< 80\%$) are kept white, while light blue areas indicate areas with sufficient coverage for the model but without any observational evidence to allow for calculation of the cost function.



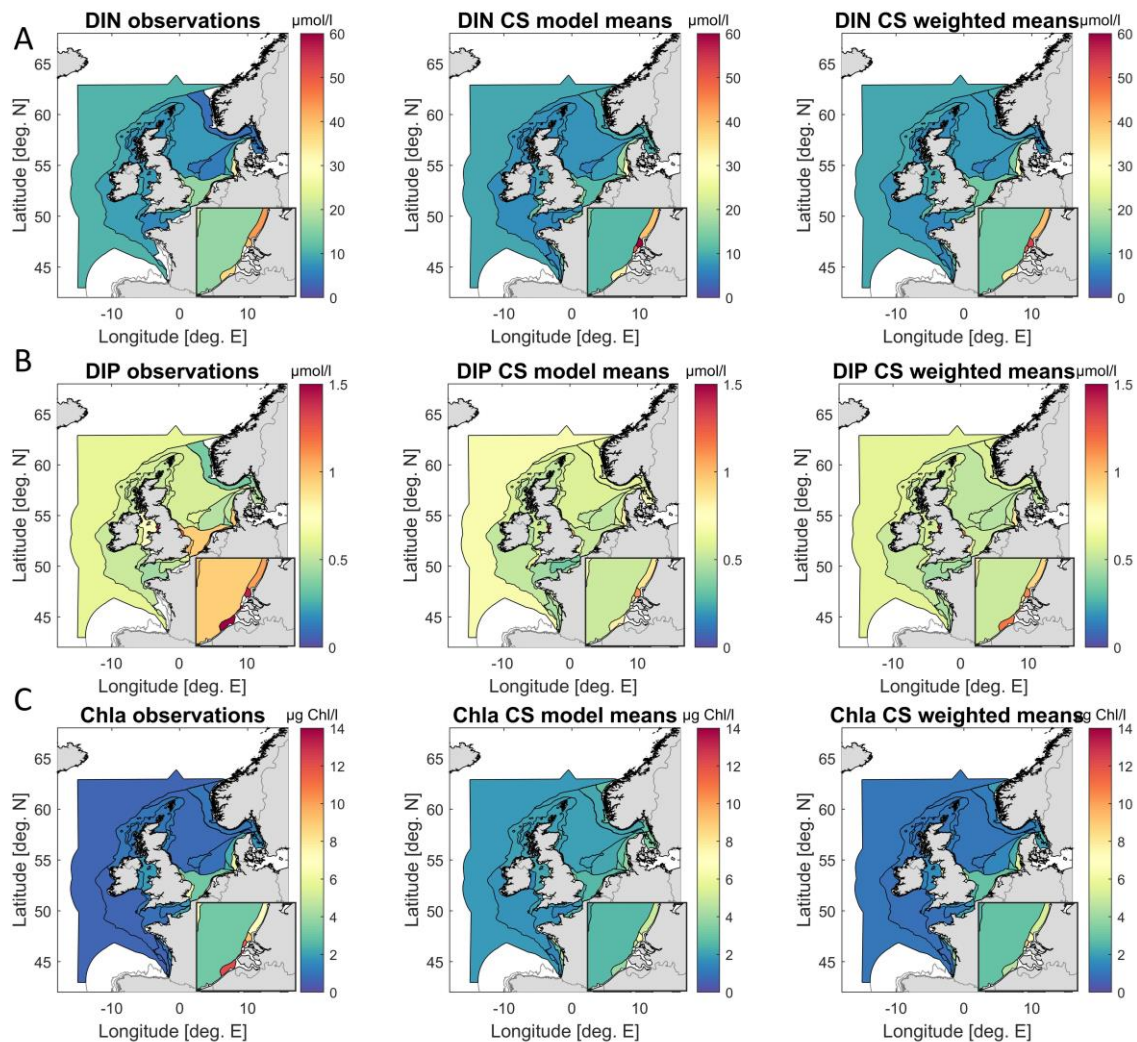
Supplementary Figure 6: DIP cost function results from the Almroth & Skogen (2010) weighting method for all models. The thick lines indicate the domain boundaries for each participating model. Areas without sufficient model coverage ($< 80\%$) are kept white, while light blue areas indicate areas with sufficient model coverage but without any observational evidence.



Supplementary Figure 7: Chl cost function results from the Almroth & Skogen (2010) weighting method for all models. The thick lines indicate the domain boundaries for each participating model. Areas without sufficient model coverage ($< 80\%$) are kept white, while light blue areas indicate areas with sufficient model coverage but without any observational evidence.

Appendix G: Example application of weighting method

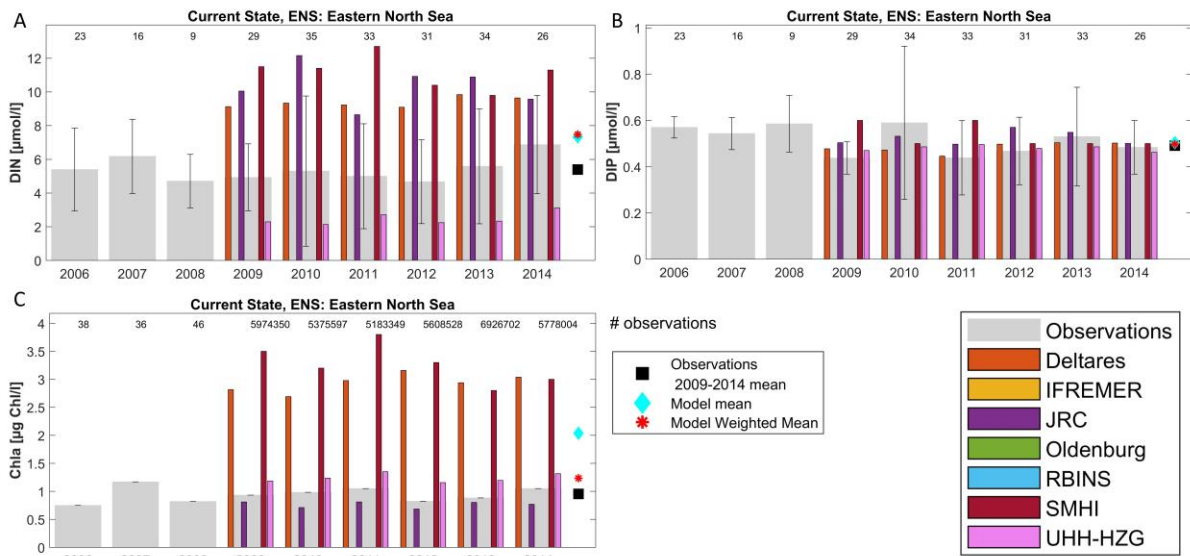
Supplementary Figure 9 shows the effect of the applied weighting method, by spatially showing the observational values (left), followed by the unweighted ensemble model mean values (center) and the weighted ensemble model mean values (right). Although the model ensemble values do not always align with the observational evidence, this may be due to the scarcity of observations in time and space in some areas.



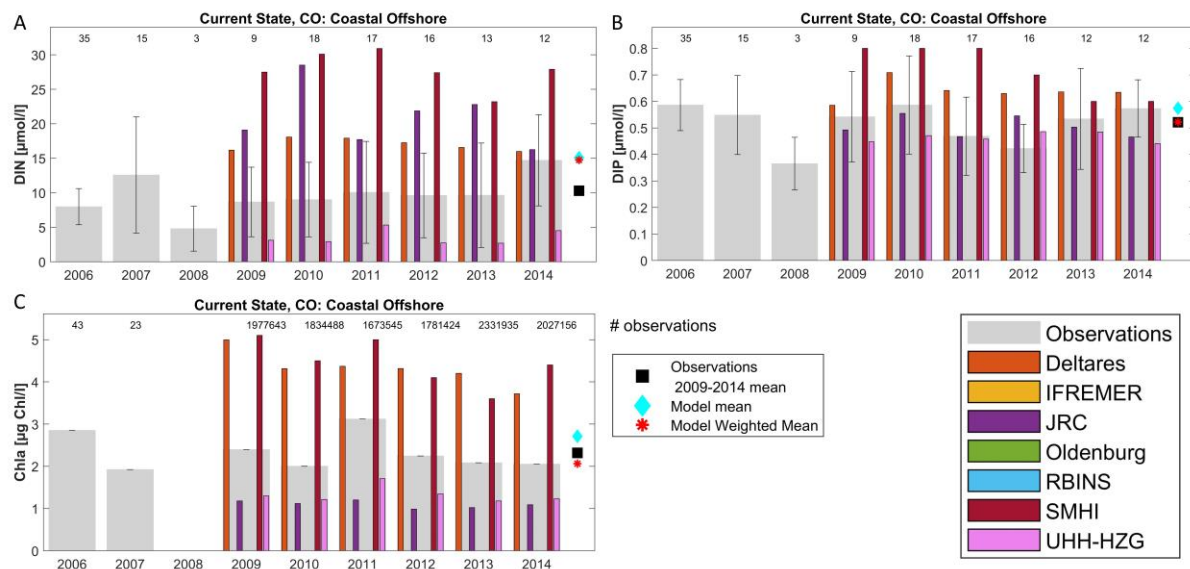
Supplementary Figure 8: spatial distributions averaged per area and over the period 2009-2014. For A) DIN, B) DIP and C) Chl showing the observational values (left), unweighted ensemble model mean (center) and weighted ensemble model mean (right).

Appendix H: Individual results for selected areas

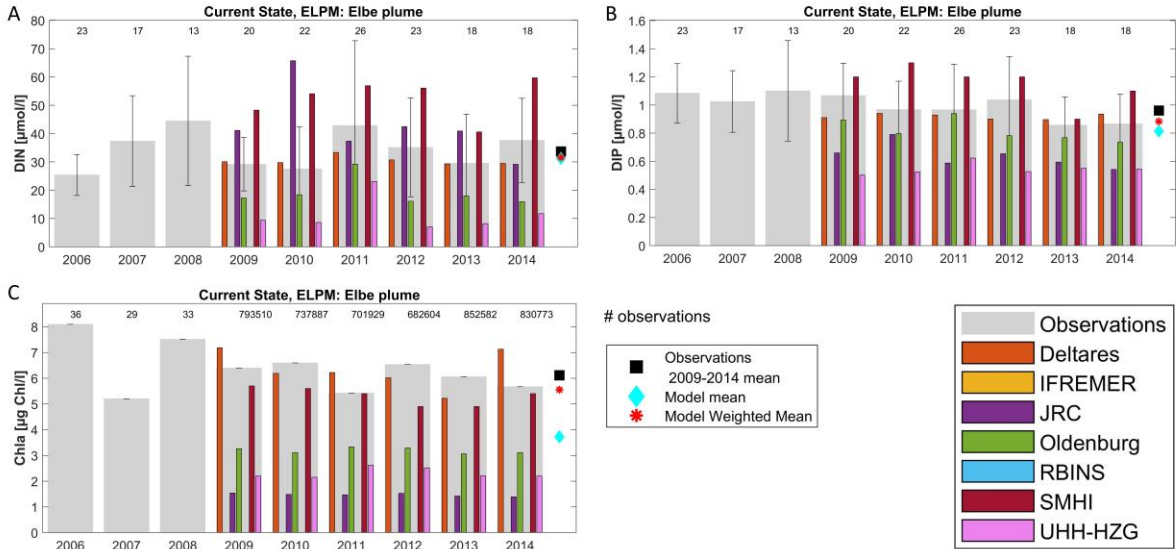
Here we show more details of the application of the weighting procedure to particular areas: the Eastern North Sea (ENS, Supplementary Figure 9), Coastal Offshore (CO, Supplementary Figure 10), Elbe plume (ELPM, Supplementary Figure 11), Rhine plume (RHPM, Supplementary Figure 12), Thames plume (THPM, Supplementary Figure 13), Irish Sea (IRS, Supplementary Figure 14) and Seine plume (SPM, Supplementary Figure 15).



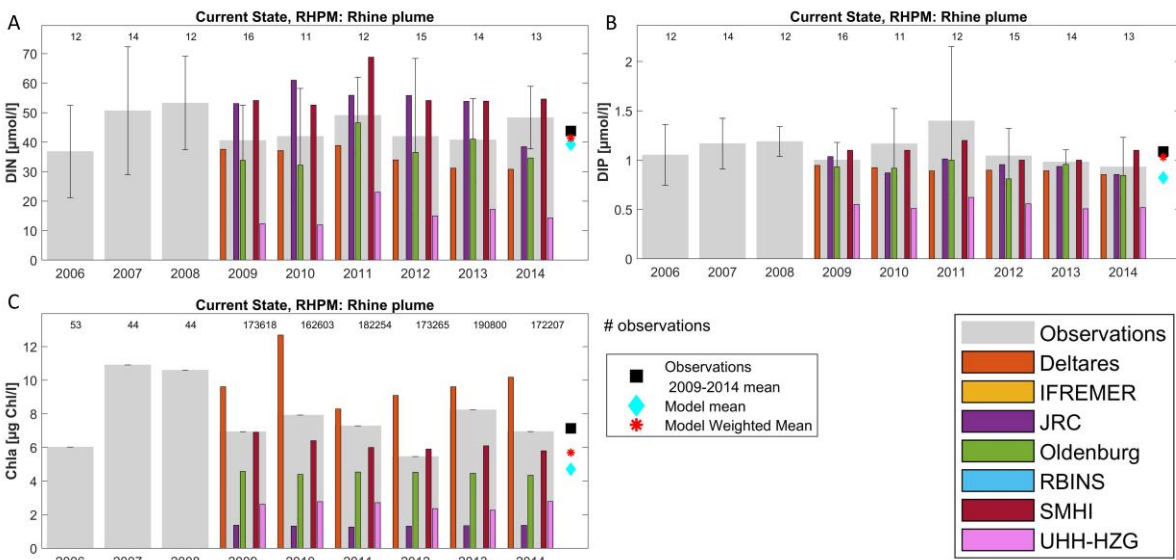
Supplementary Figure 9: annual results per model for area Eastern North Sea (ENS, 31): A) DIN, B) DIP and C) Chl. The grey bars denote the observational values per year.



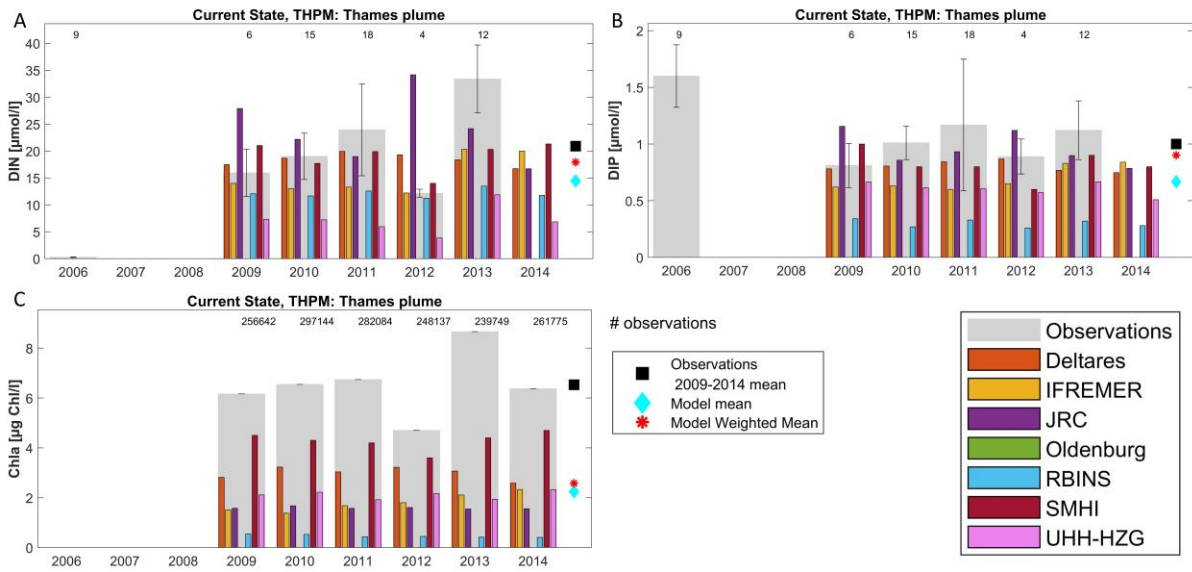
Supplementary Figure 10: annual results per model for area Coastal Offshore (CO, 30): A) DIN, B) DIP and C) Chl. The grey bars denote the observational values per year.



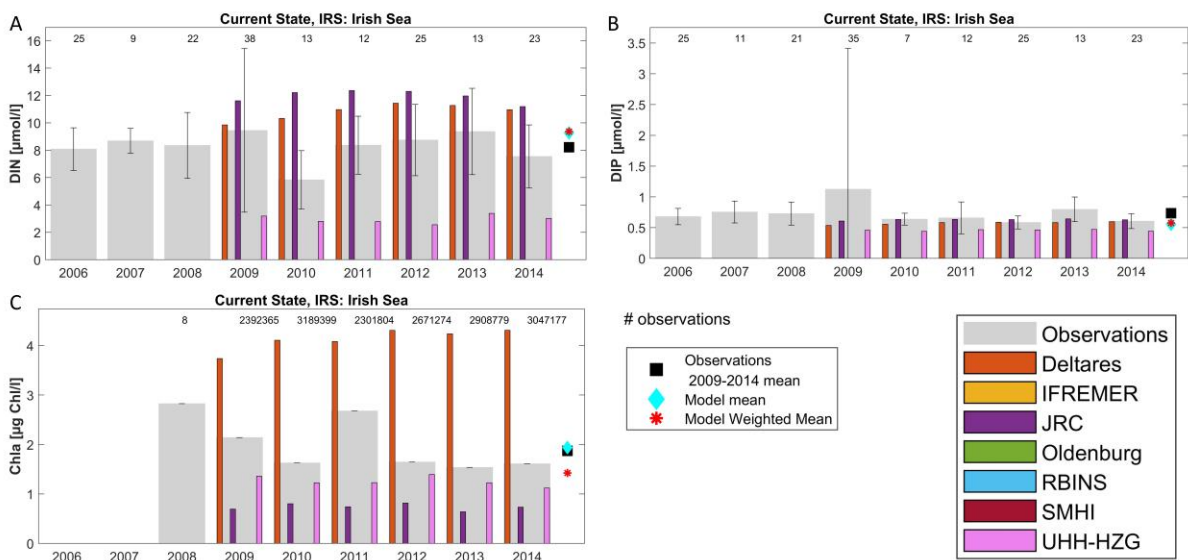
Supplementary Figure 11: annual results per model for area Elbe Plume (ELPM, 20): A) DIN, B) DIP and C) Chl. The grey bars denote the observational values per year, including their standard deviation.



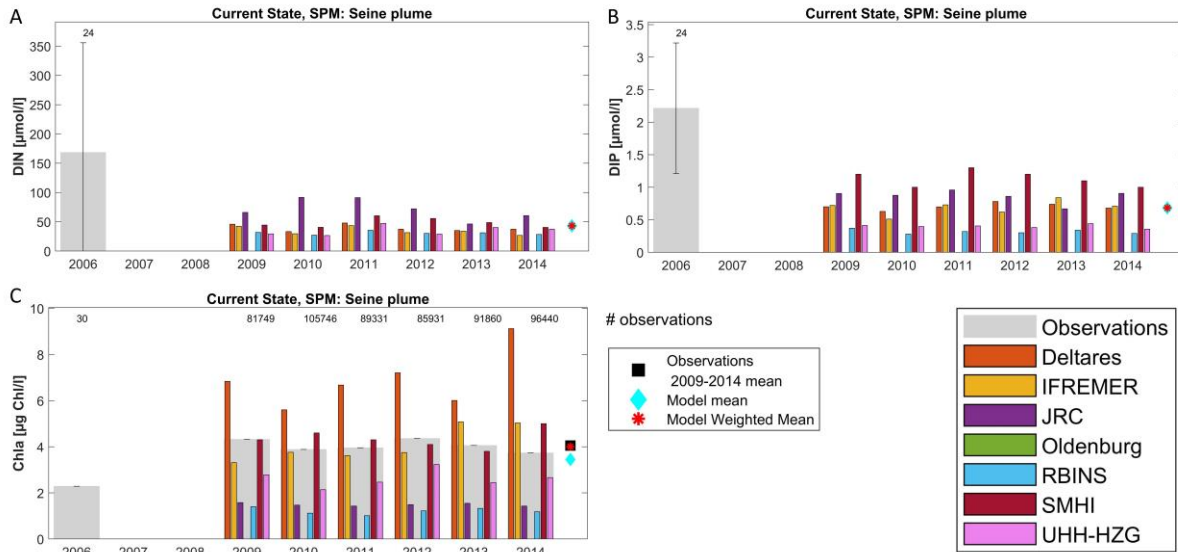
Supplementary Figure 12: annual results per model for area Rhine Plume (RHPM, 23): A) DIN, B) DIP and C) Chl. The grey bars denote the observational values per year.



Supplementary Figure 13: annual results per model for area Thames Plume (THPM, 25): A) DIN, B) DIP and C) Chl. The grey bars denote the observational values per year.



Supplementary Figure 14: annual results per model for area Irish Sea (IRS, 41): A) DIN, B) DIP and C) Chl. The grey bars denote the observational values per year.



Supplementary Figure 15: annual results per model for area Seine Plume (SPM, 15): A) DIN, B) DIP and C) Chl. The grey bars denote the observational values per year.

Appendix I: Pre-eutrophic state ensemble values

Here we include the weighted ensemble concentrations for winter DIN, DIP, total N, total P and growing season Chl and Chl P90 for all areas for the pre-eutrophic state. These values can serve as a basis for marine management policies combatting eutrophication problems. Concentrations in red denote areas and compounds where the unweighted ensemble results is used, due to a lack of observations. As Chl *in-situ* measurements were sparse, good observational coverage for Chl was obtained by inclusion of satellite data.

Supplementary Table 3: weighted ensemble mean concentrations for each area. Concentrations in red indicate areas and compounds where no observations were available, displaying the unweighted ensemble mean.

ICG EMO ID	Pre-eutrophic state Area	DIN	DIP	TotalN	TotalP	Chl	Chl P90
		µmol/l	µmol/l	µmol/l	µmol/l	µg Chl/l	µg Chl/l
1	CFR	10.53	0.40	9.80	0.41	1.88	3.00
2	CCTI	8.01	0.43	11.12	0.45	1.56	2.51
3	ATL	10.23	0.65	13.89	0.96	1.22	2.80
4	SHPM	7.38	0.53	8.74	0.61	1.22	2.49
5	CNOR1	8.34	0.58	9.00	0.64	1.83	4.25
6	CNOR2	6.89	0.51	7.58	0.55	1.29	2.69
7	CNOR3	6.10	0.45	6.83	0.57	1.57	2.96
8	DB	4.80	0.50	4.39	0.46	0.88	1.94
9	KD	4.43	0.46	5.17	0.54	1.84	3.39
10	NT	7.28	0.58	9.51	0.72	1.12	2.77
11	SNS	8.65	0.47	6.64	0.46	2.51	3.24
12	GBC	4.83	0.46	5.07	0.51	1.79	2.86
13	ADPM	5.91	0.45	6.24	0.48	1.15	2.47

14	GBSW	5.80	0.46	6.25	0.49	0.58	1.47
15	SPM	25.25	0.61	19.85	0.54	3.41	4.79
16	GDPM	8.49	0.45	10.42	0.62	3.61	6.64
17	CUKC	8.53	0.49	9.19	0.50	1.51	2.60
18	CWMTI	6.14	0.46	7.50	0.47	0.97	1.80
19	SCHPM1	17.13	0.77	18.24	0.83	3.32	4.73
20	ELPM	17.37	0.53	14.04	0.48	3.50	4.87
21	SCHPM2	22.06	0.57	23.68	0.72	5.56	9.20
22	MPM	27.08	0.76	24.30	0.67	4.47	7.05
23	RHPM	19.71	0.66	15.86	0.26	4.14	4.42
24	EMPM	7.20	0.45	7.36	0.58	3.53	4.63
25	THPM	9.69	0.60	8.32	0.67	2.11	3.16
26	HPM	17.42	0.77	13.23	0.68	4.92	6.72
27	ECPM1	7.31	0.52	7.60	0.50	1.39	3.35
28	ECPM2	7.25	0.57	3.23	0.50	2.35	3.80
29	IS2	7.51	0.58	8.14	0.63	1.16	3.09
30	CO	8.89	0.44	7.71	0.52	1.82	3.31
31	ENS	5.38	0.45	5.36	0.43	1.15	2.57
37	ASS	7.77	0.56	8.43	0.61	0.92	1.96
38	CIRL	7.57	0.51	7.82	0.54	1.20	2.59
39	CUK1	7.80	0.55	8.82	0.60	1.14	2.44
40	IS1	9.14	0.60	9.48	0.65	1.10	2.70
41	IRS	6.61	0.52	7.32	0.57	1.31	2.77
42	KC	5.03	0.43	5.07	0.47	1.58	3.14
43	NNS	7.64	0.59	7.86	0.41	1.05	2.39
44	CWM	5.50	0.44	6.12	0.48	0.87	1.87
45	LBPM	19.54	0.91	19.30	0.92	4.91	8.88
46	SK	4.47	0.47	5.97	0.55	1.15	2.45
47	SS	6.44	0.53	7.01	0.57	0.98	2.50
52	LPM	12.89	0.53	12.15	0.62	2.23	3.81
53	GBCW	7.86	0.50	8.65	0.56	1.78	3.75

References

Almroth-Rosell, E., Eilola, K., Hordoir, R., Meier, H.M. and Hall, P.O., 2011. Transport of fresh and resuspended particulate organic material in the Baltic Sea—a model study. *Journal of Marine Systems*, 87(1), pp.1-12

Almroth, E. and Skogen, M.D., 2010. A North Sea and Baltic Sea model ensemble eutrophication assessment. *Ambio*, 39(1), pp.59-69.

Baretta, J.W., Ebenhöf, W. and Ruardij, P., 1995. The European regional seas ecosystem model, a complex marine ecosystem model. *Netherlands Journal of Sea Research*, 33(3-4), pp.233-246. Blauw, A.N., Los, H.F., Bokhorst, M. and Erfteimeijer, P.L., 2009. GEM: a generic ecological model for estuaries and coastal waters. *Hydrobiologia*, 618(1), pp.175-198.

Dulière V., Gypens N., Lancelot C., Luyten P., Lacroix G. (2019) Origin of nitrogen in the English Channel and Southern Bight of the North Sea ecosystems. *Hydrobiologia* 845: 13-33 – Doi: [10.1007/s10750-017-3419-5](https://doi.org/10.1007/s10750-017-3419-5)

Eilola, K., Meier, H.M. and Almroth, E., 2009. On the dynamics of oxygen, phosphorus and cyanobacteria in the Baltic Sea; A model study. *Journal of Marine Systems*, 75(1-2), pp.163-184

Enserink, L., Blauw, A., van der Zande, D. and Markager, S. (2019). Summary report of the EU project ‘Joint monitoring programme of the eutrophication of the North Sea with satellite data’ (Ref: DG ENV/MSFD Second Cycle/2016). Belgium: REMSEM, 21

Friedland, R., Stips, A., Grizzetti, B., de Roo, A. and Lessin, G., 2020. Report on the biogeochemical model of the North-Western European Shelf. *Publications Of-fice of the European Union, Luxembourg, ISBN*, pp.978-92.

Große, F., Kreuz, M., Lenhart, H.J., Pätsch, J. and Pohlmann, T. (2017) A novel modeling approach to quantify the influence of nitrogen inputs on the oxygen dynamics of the North Sea. *Frontiers in Marine Science*, 4, p.383

Heath, M. R., Edwards, A. C., Patsch, J., & Turrell, W. R. (2002). Modelling the behaviour of nutrients in the coastal waters of Scotland. Report. <https://strathprints.strath.ac.uk/18568/6/strathprints018568.pdf>

Hordoir, R., Axell, L., Höglund, A., Dieterich, C., Fransner, F., Gröger, M., Liu, Y., Pemberton, P., Schimanke, S., Andersson, H. and Ljungemyr, P. (2019) Nemo-Nordic 1.0: a NEMO-based ocean model for the Baltic and North seas—research and operational applications. *Geoscientific Model Development*, 12(1), pp.363-386.

ICES (2022) ICES Data Portal, Dataset on Ocean HydroChemistry, Extracted April 21, 2022. ICES, Copenhagen

Kerimoglu, O., Voynova, Y.G., Chegini, F., Brix, H., Callies, U., Hofmeister, R., Klingbeil, K., Schrum, C. and van Beusekom, J.E. (2020) Interactive impacts of meteorological and hydrological conditions on the physical and biogeochemical structure of a coastal system. *Biogeosciences*, 17(20), pp.5097-5127

Lancelot C, Spitz Y, Gypens N, Ruddick K, Becquevort S, Rousseau V, Lacroix G, Billen G. (2005) Modelling diatom and *Phaeocystis* blooms and nutrient cycles in the Southern Bight of the North Sea: the MIRO model. *Marine Ecology Progress Series* 289 : 63-78 – Doi: [10.3354/meps289063](https://doi.org/10.3354/meps289063)

Lazure, P. and Dumas, F. (2008) An external–internal mode coupling for a 3D hydrodynamical model for applications at regional scale (MARS). *Advances in water resources*, 31(2), pp.233-250.

Luyten P. (2011) COHERENS – A Coupled HydrodynamicalEcological Model for Regional and Shelf Seas: User Documentation. Version 2.0. RBINS-MUMM Report, Royal Belgian Institute of Natural Sciences.

Lyard, F., Lefevre, F., Letellier, T. et al. (2006) Modelling the global ocean tides: modern insights from FES2004. *Ocean Dynamics* 56, 394–415.

Madec, G., 2010. Nemo ocean engine, v. 3.3. *IPSL Paris France*.

Menesguen, A., Lacroix, G. (2018) Modelling the marine eutrophication: A review. *Science of the Total Environment* 636: 339-354. <https://doi.org/10.1016/j.scitotenv.2018.04.183>

Ménesguen, A., Dussauze, M., Dumas, F., Thouvenin, B., Garnier, V., Lecornu, F. and Répécaud, M., (2019) Ecological model of the Bay of Biscay and English Channel shelf for environmental status assessment part 1: Nutrients, phytoplankton and oxygen. *Ocean Modelling*, 133, pp.56-78

Ruardij, P. and Van Raaphorst, W., 1995. Benthic nutrient regeneration in the ERSEM ecosystem model of the North Sea. *Netherlands Journal of Sea Research*, 33(3-4), pp.453-483.

Sirjacobs, Damien, et al. (2011) Cloud filling of ocean colour and sea surface temperature remote sensing products over the Southern North Sea by the Data Interpolating Empirical Orthogonal Functions methodology. *Journal of Sea Research* 65.1: 114-130.

Troost, T.A., Desclaux, T., Leslie, H.A., van der Meulen, M.D. and Vethaak, A.D. (2018) Do microplastics affect marine ecosystem productivity?. *Marine pollution bulletin*, 135, pp.17-29.

Troost, T.A., Wijsman, J.W.M., Saraiva, S. and Freitas, V. (2010) Modelling shellfish growth with dynamic energy budget models: an application for cockles and mussels in the Oosterschelde (southwest Netherlands). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1557), pp.3567-3577.

van der Molen, J., Aldridge, J.N., Coughlan, C., Parker, E.R., Stephens, D. and Ruardij, P. (2013) Modelling marine ecosystem response to climate change and trawling in the North Sea. *Biogeochemistry*, 113(1), pp.213-236

van der Molen, J., Ruardij, P., Greenwood, N. (2017) A 3D SPM model for biogeochemical modelling, with application to the northwest European continental shelf, *J of Sea Res.*, 127, 63:81, DOI: 10.1016/j.seares.2016.12.003

van Leeuwen, S., Tett, P., Mills, D. and van der Molen, J. (2015). Stratified and nonstratified areas in the North Sea: Long-term variability and biological and policy implications. *Journal of Geophysical Research: Oceans* 120: 4670-4686

Vichi, M., Pinardi, N. and Masina, S. (2007). A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: Theory. *Journal of Marine Systems*, 64(1-4), pp.89-109

Zijl, F. , S.C. Laan, J. Groenenboom, (2021). Development of a 3D model for the NW European Shelf (3D DCSM-FM). Deltares report 11205259-015-ZKS-003, 61 pp. https://publications.deltares.nl/11205259_015.pdf