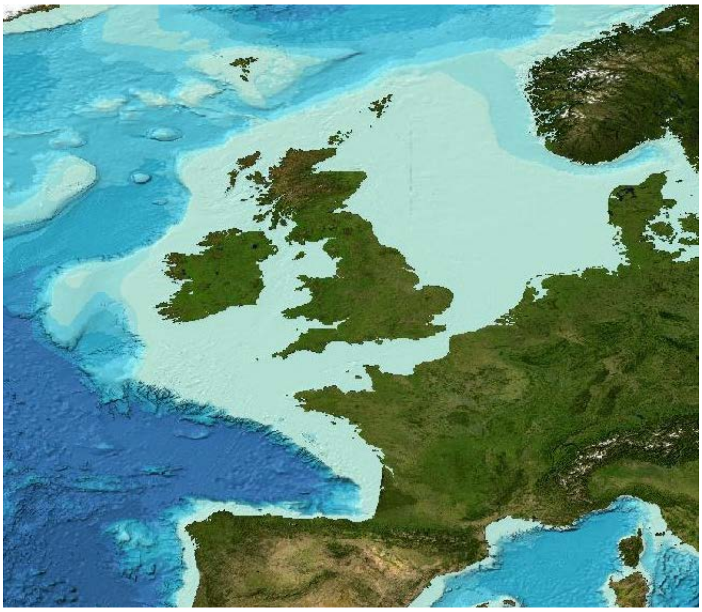
Assessment of model skill in the North East Atlantic

by the modelling groups at PML (Plymouth Marine Laboratory) and Cefas, UK

Contents

Summary **2**
Introduction **2**
North East Atlantic **2**
1. POLCCOMS-ERSEM 2
1.1 Description 2
1.2 Set-up of the hindcast simulation. 3
1.3 Model Hindcast 4
1.4 The in situ data 7
1.5 Validation procedure 8
1.6 Results 9
1.7 Summary of hindcast experiment and lessons learnt 12
1.8 References 16
2. Ecospace driven by POLCOMS-ERSEM 16
2.1 Description NE Atlantic 16
2.2 Summary and lessons learnt 19
2.3 Size-based modelling of higher trophic levels (Cefas, PML) 20
2.4 References 22



Summary

Regional hindcast simulations for the period 1990-2009 have been made using the OPEC regional model systems for both the biogeochemical and HTL model components. This document records the assessment of the model system's performance. The hindcast skill is evaluated using the benchmark tool developed in the framework of OPEC.

The results indicate that all regional model systems demonstrate a range of skill, depending on the variables chosen. Physical variables (e.g. T, S) are generally have the most skill) followed by chemical variables (e.g. O_2, Nutrients) then plankton variables (e.g. chlorophyll) for the coupled hydrodynamic LTL models. The HTL models have more skill for small pelagic fish (e.g. Sprat in the Baltic and Anchovy in the Aegean) than the plankton model that drives them. However the skill for larger pelagic and demersal fish is generally poor.

Introduction

A primary objective of OPEC was to set-up the ecological model system for the next generation GMES marine ecological service in European Seas. Each regional model system comprises a core coupled hydrodynamic-plankton model, a HTL component, a representation of the carbon chemistry and a data assimilation system. These have been used to perform 20yr hindcast of each region and to benchmark model performance. This document focuses on assessing the performance of the hindcast model systems with an emphasis on evaluating the performance of key policy relevant metrics. The goal is to benchmark the quality of the regional hindcasts, firstly to inform users of the data products about the skill and hence usefulness of the simulations, secondly to provide a benchmark against which future model development can be assessed.

Model skill scores are colour coded to give an indication of model skill as follows; Good = green, Moderate = Black, Poor = Red; correlation r Good > 0.75, poor < 0.20, M good = > 0.75, Poor = < 0, Reliability index Good 0.8-1.2, Poor > 2 or –ve. PBias, Good = < 20%, Poor = > 100%.

North East Atlantic

1. POLCCOMS-ERSEM

1.1 Description

The North East Atlantic region is represented in figure 1. The model deployed for the hindcast of the hydrodynamic and lower trophic level ecosystem component in this region is the coupled POLCOMS-ERSEM system.

This is a well-established system subject to numerous peer-reviewed scientific publications (Allen et al. 2001; Siddorn et al. 2007; Holt et al. 2012) and applied in a variety of international research projects covering subjects from operational forecast over climate change projections to impact studies, such as ocean acidification or fisheries management. This system was preferred over the next generation operational model due to its robustness and efficiency, two elements that are crucial for the data assimilation component within this project involving the computational heavy Ensemble Kalman Filter.

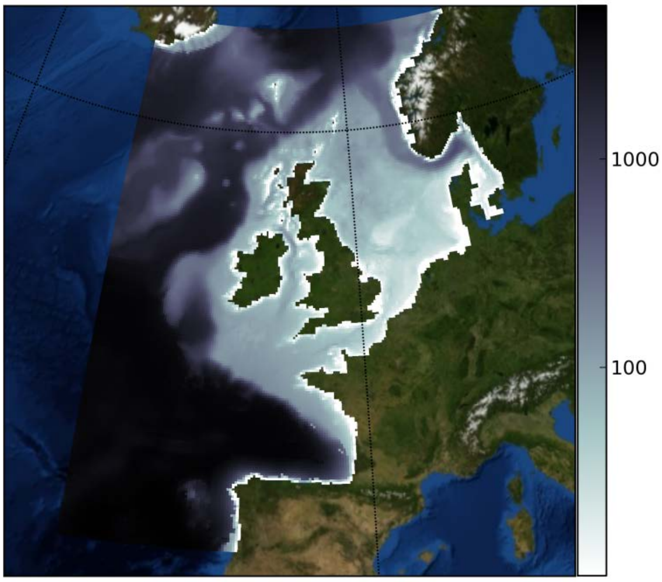


Figure 1. The North East Atlantic region.

1.2 Set-up of the hindcast simulation.

The model system of the NE Atlantic has been applied in a nineteen year-long hindcast simulation for the years 1991-2009. The simulation was started using the December 1990 output for the same domain from a previous project (EC FP7 MEECE). The model was spun-up by running repeatedly year 1991 for 3 times, using OPEC forcings for that year. The hindcast was then initialised using the last conditions obtained for January 1991.

In the hindcast simulation, the atmospheric forcing was given by data from the DMI regional climate model described in a previous deliverable of this Project (Deliverable D2.2: "Meta data for the boundary conditions and forcing functions for each region"). The data used for the NE Atlantic model configuration are summarized in Table 1.1, for the sake of commodity.

Table 1.1. List of atmospheric surface variables used for the NE Atlantic model system.

Name	code	Level variable	Description	Unit	Resolution
MSLP	1	0m	Mean sea level pressure	Pa	Every 3 h
TEMP2	11	2m	2-m temperature	K	Hourly
U10	33	10m	10-m zonal wind	m/s	Hourly
V10	34	10m	10-m meridional wind	m/s	Hourly
Q2M	51	2m	2-m specific humidity	Kg/kg	Every 3h
ACLCOV	71	0m	Total cloud cover	-	Daily

APRS	65	0m	Snow fall	Mm/d	Daily
APRL	62	0m	Large scale precipitation	Mm/h	Hourly
APRC	63	0m	Convective precipitation	Mm/h	Hourly
SRADS	111	0m	Net SW radiation at surface	W/m2	Every 3h
TRADSU	115	0m	Downw./Upw. LW radiation at surface	W/m2	Every 3h

The description of the general configuration of the oceanic and riverine boundary of the North East Atlantic model system is given in the deliverable D2.4 "Description of the coupled model for each region". In the hindcast simulation, the oceanic conditions at the open boundaries (temperature, salinity, currents and sea surface elevation) were extracted for the years 1991-2009 from the GLORYS reanalysis product (MERCATOR) provided within the MyOcean project. The corresponding conditions for dissolved nutrients and oxygen were extracted from WOA climatology (2005) dissolved inorganic carbon from GLODAP. Other ecosystem variables were treated applying no-flux conditions at the open boundaries.

For freshwater fluxes, daily discharge data for 250 rivers were used from the Global River Discharge Data Base and from data prepared by the Centre for Ecology and Hydrology. River nutrient loading matches that used by Lenhart et al. (2010), with raw data for the UK, Northern Ireland, Ireland, France, Norway, Denmark and the Baltic processed by van Leeuwen (CEFAS, UK) and raw data for Germany and the Netherlands was processed by Pätsch and Lenhart (2004). In addition, Baltic inflow at the belt was represented as river-inflow. Atmospheric input of nutrients was derived from EMEP.

1.3 Model Hindcast

The hindcast results for the validation variables are summarized in figure 2. The maps show the average annual distributions in the years 1991-2009. Surface values are represented in the figures.

Based on the simulations, which are consistent with those obtained in previous studies in the NE Atlantic region, (Holt 2012, Artioli 2012) the domain can be roughly classified into three regimes:

- A shallow estuarine or coastal regime extending along the continental North Sea coast and around the British Isles with high productivity (as indicated by the chlorophyll-a distribution), often under strong riverine influences (see nutrient and salinity fields). The balance of nitrate and phosphate in these areas clearly reflects the generally phosphorus limited growth condition in these waters.
- A shelf-seas regime extending over the off-shore areas of the NW-European shelf seas with intermediate production levels and nutrient supply and lower pH values than the off-shelve areas.
- The off-shelf regions that are further subdivided by the gulf-stream in a Southern part under typical open ocean conditions with low nutrient supply and low ventilation and consequently low production and a Northern area under sub polar influence with high oxygen levels and elevated nutrient supply but low temperature and light conditions limiting growth of biomass.

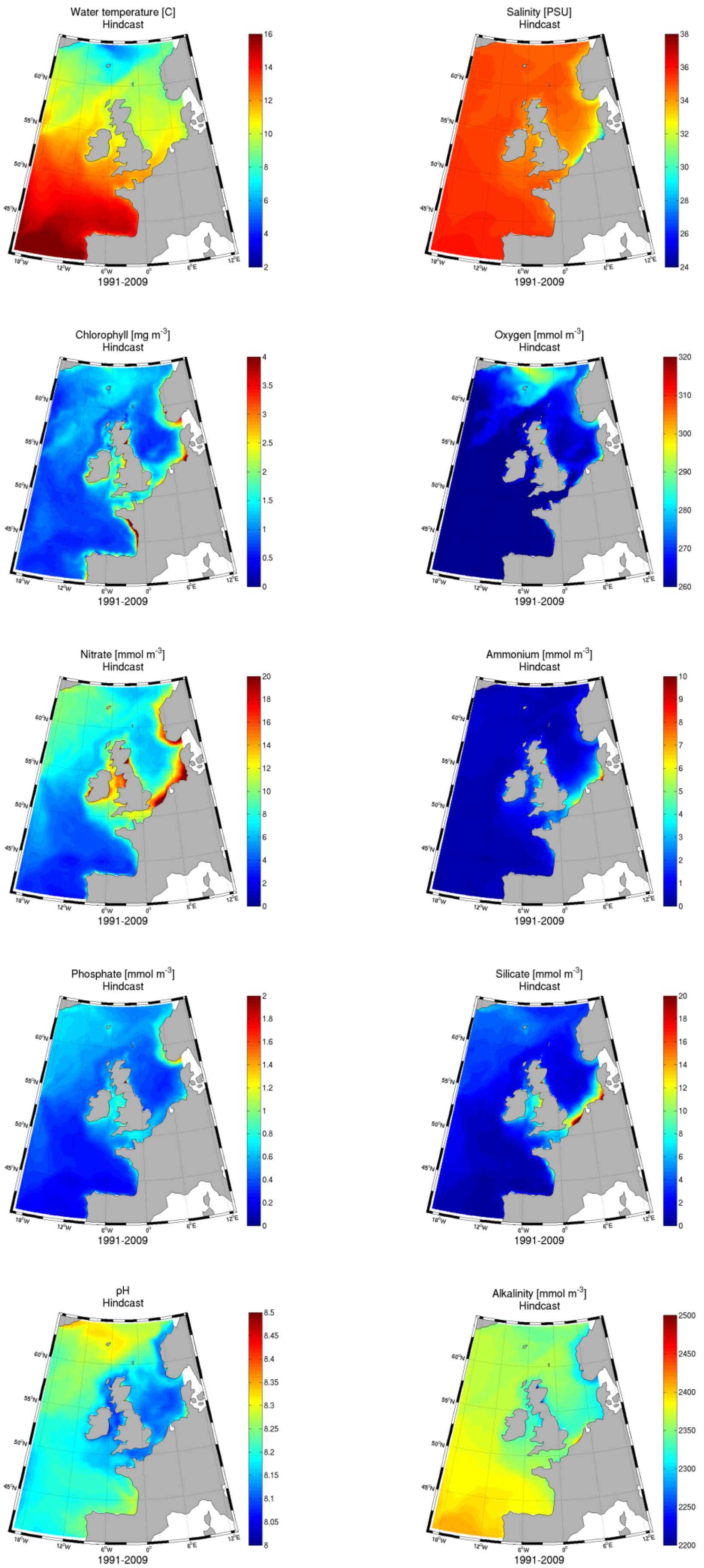
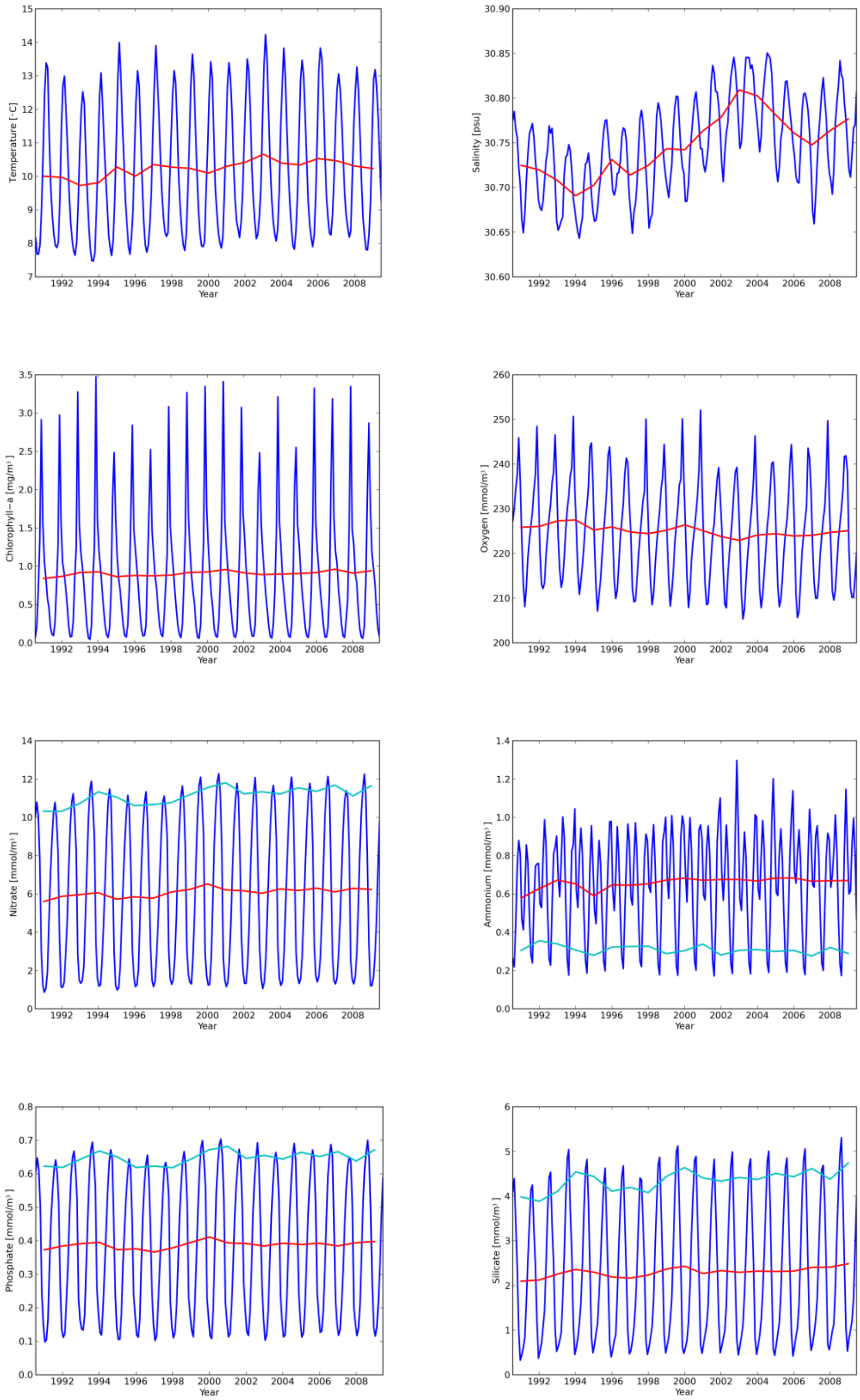

Water temperature [C]
Hindcast
1991-2009
Salinity [PSU]
Hindcast
1991-2009
Chlorophyll [mg m-3]
Hindcast
1991-2009
Oxygen [mmol m-3]
Hindcast
1991-2009
Nitrate [mmol m-3]
Hindcast
1991-2009
Ammonium [mmol m-3]
Hindcast
1991-2009
Phosphate [mmol m-3]
Hindcast
1991-2009
Silicate [mmol m-3]
Hindcast
1991-2009
pH
Hindcast
1991-2009
Alkalinity [mmol m-3]
Hindcast
1991-2009

Figure 2. Hindcast results: average annual surface distributions of the validation variables in the years 1991-2009.



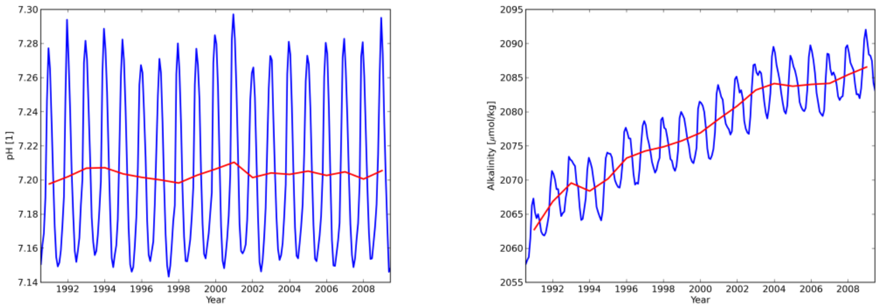


Figure 3. Hindcast results: time series of monthly (blue) and annual (red) domain averages of surface variables for the years 1991-2009. Winter nutrient averages are shown in light blue.

Time-series plots for the same indicators are presented in figure 3. It can be seen that for most of the variables the interannual variability of the domain averages is comparatively weak and dominated by the seasonal cycle with no significant trends visible over the simulated period with the exception of increasing salinity and alkalinity. Temperature, nutrients and chlorophyll show some tendency to increase while oxygen slightly decreases, but all of these signals are too weak to indicate a general trend in the ecosystem.

1.4 The in situ data

The *in situ* data used for the validation of the hindcast simulation were taken from the International Council for the Exploration of the Sea (ICES) Ecosystem Data Online Warehouse (ICES, 2009). Ten variables were considered (Table 1.2), ranging from chlorophyll to nutrient concentrations, from physical to carbonate system data.

Table 1.2 Name, unit and representative symbols of the variables considered in the hindcast validation.

Variable	Unit	Model symbol
Water temperature	C	ETW
Salinity	PSU	x1X
chlorophyll	mg m^{-3}	Chl
Dissolved oxygen	mmol m^{-3}	O2o
Nitrate	mmol m^{-3}	N3n
Ammonia	mmol m^{-3}	N4n
Phosphate	mmol m^{-3}	N1p
Silicate	mmol m^{-3}	N5s
pH		pH
Alkalinity	mmol m^{-3}	Alk

The initial data set included all the ICES bottle and CTD data available in the region and depths covered by the model domain, and spanning the years 1991-2009. The data were acquired in a comma-separated- variable format (.csv).

The initial dataset underwent a preliminary quality control (QC) processing, by means of a Python tool developed *ad hoc.*

This processing followed the QC procedure applied by De Mora et al. (2013) with ICES data. In particular:

- we removed repeated data, i.e. data with identical measurement time, longitude, latitude, depth and value;
- all chlorophyll measurements below 100 m were removed from the dataset , since we found a large proportion of measurements with a value of exactly 0.1, even at depth below 1km, as De Mora et al.,2013).
- data values minor or equal than zero were removed for all the variables, but temperature, because interpreted as "no data" entries.
- data with missing value code were removed;
- measurement time was arbitrarily set equal to 12:00, when the time was not reported in the database. We note that measurement time is not relevant in our validation since data were compared with monthly means;
- the measurement unit of alkalinity data was changed from the original mmol/l to mmol/m3, which is consistent with ERSEM output.

The database obtained from the preliminary QC was applied in the validation of the hindcast according to the procedure described in the next section.

1.5 Validation procedure

The hindcast simulation was benchmarked by comparing the monthly averages with the ICES in situ data, for the variables listed in Table1.2.

The validation was carried out by using the OPEC benchmarking tool (see Deliverable 2.3 "Target variables and benchmarking metrics: Technical Specification of Benchmarking Tool v1").

In our application, the data-to-model match-ups between model output and data were obtained by setting maximal temporal distances equal to fifteen days and vertical spatial steps equal to 2.5 m. The choice of the temporal distance derives from our use of monthly outputs in the benchmarking analysis. The choice for the vertical spatial distance follows the re-gridding of the model and data values required by the application of the OPEC tool. In our application, the values were re-gridded with vertical steps of 5m in the first 50 m of the water column, thus we choose half the step (2.5 m) for the match-up vertical distance.

In the benchmarking analysis, we discarded data-to-model match-ups that were close to the geographical limits of the model domain (i.e. within 0.5° from the limits, in latitude or longitude), to avoid misleading effects of the boundary conditions. Furthermore, we excluded the match-ups from the Kagerrak/Kattegat area (i.e. the Danish area east of 9° Longitude), where the model is affected negatively by the Baltic boundary condition (see e.g. Artioli et al., 2012).

Finally we discarded match-ups where model or data values were outliers. The thresholds for high-value outliers were defined as the 95th percentiles of the distributions. The threshold for low-value outliers were set on the base of regional studies and expert judgement (Artioli et al., 2012; MEECE, 2013). The thresholds are listed separately for model and data values in Table 1.3

Table 1.3. The table lists the lower and upper threshold values we applied to define the outliers in the model and in situ data. Upper values correspond to the 95th percentiles of the data or model distributions.

Variable	Unit	Model lower	Model upper	Data lower	Data upper
Water temperature	C	-2.	14.58	-2.	15.94
Salinity	PSU	30.	35.45	30.	35.6
Chlorophyll	mg m^{-3}	1E-7	4.8	1E-7	10.27
Dissolved oxygen	mmol m^{-3}	100.	300.34	100.	334.96
Nitrate	mmol m^{-3}	1E-7	29.9	1E-7	36.
Ammonia	mmol m^{-3}	1E-7	7.44	1E-7	8.3
Phosphate	mmol m^{-3}	1E-7	1.4	1E-7	1.24
Silicate	mmol m^{-3}	1E-7	18.76	1E-7	17.3
pH		7.	8.4	7.	8.46
Alkalinity	mmol m^{-3}	2200	2394.	2200	2419

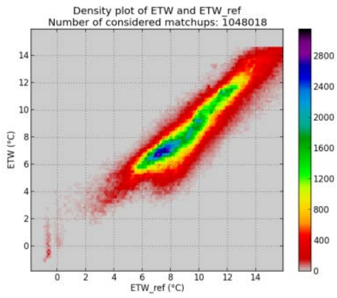
1.6 Results

A first overview of the results of the model-to-data match-up analysis are presented in figure 4, which shows the density plots for each validation variable. The number of valid match-ups resulting from the data quality control procedure and from the selection using threshold values (see section 2.5) is reported in the heading of each graph.

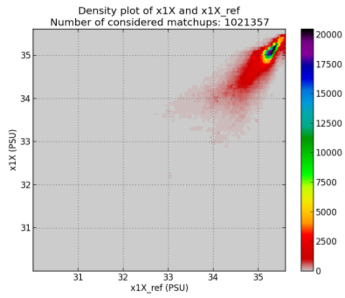
The major variables temperature, salinity, nutrients and chlorophyll-a show a good match with the bulk of the data distributed around the 45 diagonal of the diagram indicating perfect model-data match. The performance of Ammonium, pH and alkalinity is considerably weaker. In addition, for nitrate and phosphate vertical lines at low in-situ data concentrations appear indicating some irregularities the causes of which need further investigation. (Possible causes next to poor model performance are very localised phenomena not resolved by the current spatial resolution or spurious entries in the original dataset, which would require an improved quality control.)

Following this first qualitative estimate of model performance and based on the matched-up database of model and in-situ data a standardised statistical analysis was performed.

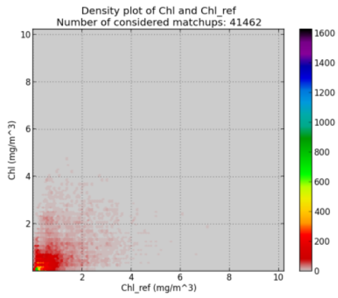
Water temperature



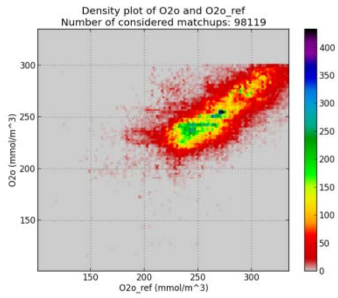
Salinity



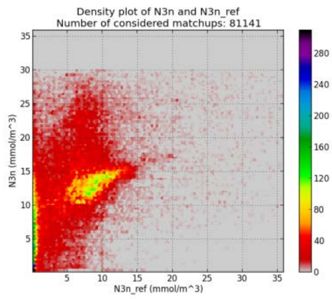
Chlorophyll



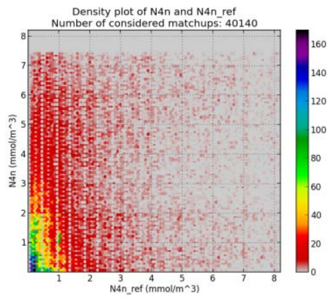
Dissolved oxygen



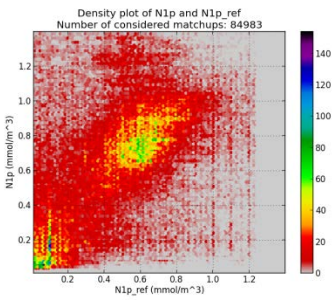
Nitrate



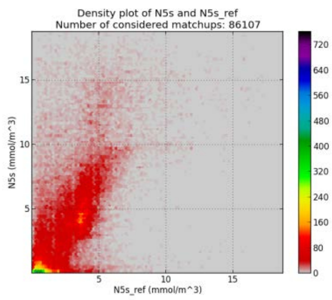
Ammonium



Phosphate



Silicate



pH

Alkalinity

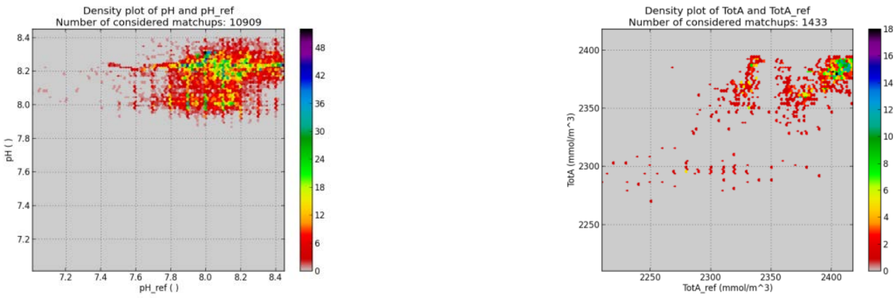


Figure 4. Density plot of the valid model-to-data match-ups for the ten validation variables.

The results of this quantitative benchmarking analysis are presented in table 2.4 and Figures 4-6. The table lists the values of core skill metrics computed by the OPEC benchmarking toolbox, for each validation variable. The Figures helps the description of the table by visualizing the values of some of the metrics, in the form of Taylor and Target diagrams.

The results indicate that the model had some skill in hindcasting the in situ data of several of the validation variables considered for the North East Atlantic region.

Table 2.1.4 Results of the hindcast benchmarking for the NE Atlantic region: Core skill metrics computed for the model-to-data matchups. Model skill scores are colour coded to give an indication of model skill as follows; Good = green, Moderate = Black, Poor = Red; correlation r good > 0.75, poor < 0.20, ME good = > 0.5, Poor = < 0, Reliability index good 0.8-1.2, poor > 2 or –ve. PBias, good = < 20%, poor = > 100%.

Region: NE Atlantic Model: POLCOMS-ERSEM Met forcing: DMI regional hindcast Hindcast time period: 1991-2009 Contact: M Butenschoen momm@pml.ac.uk							
		Hindcast					
Variables	**Unit**	**Model Efficiency**	**% Model Bias**	**Pearson Correlation Coefficient**	**RMSE**	**Unbiased RMSE**	**Reliability Index**
ETW	C	**0.73**	**9.22**	**0.91**	**1.40**	**1.10**	**1.09**
x1X	PSU	**0.60**	**0.58**	**0.82**	**0.53**	**0.49**	**1.01**
Chl	mg m^{-3}	**0.11**	**33.65**	**0.45**	**1.80**	**1.72**	**2.05**
O2o	mmol m^{-3}	**0.30**	**5.12**	**0.70**	**25.92**	**22.11**	**1.05**
N3n	mmol m^{-3}	**-1.64**	**-94.29**	**0.38**	**9.41**	**6.87**	**2.00**
N4n	mmol m^{-3}	**-1.26**	**-63.38**	**0.25**	**2.44**	**2.26**	**2.09**
N1p	mmol m^{-3}	**-0.67**	**-37.15**	**0.47**	**0.36**	**0.32**	**1.50**
N5s	mmol m^{-3}	**-1.35**	**-55.64**	**0.55**	**4.12**	**3.66**	**1.74**

pH		-0.35	-1.21	0.22	0.22	0.19	1.01
Alk	mmol m^{-3}	0.40	0.31	0.67	29.03	28.10	1.01

Table and figures suggests that a good hindcast was achieved for water temperature, salinity, chlorophyll, dissolved oxygen and alkalinity data. For all this variables, the model efficiency values were positive (Table 2.4), indicating that the model errors were lower than the variability of the data. The points representing these variables in the Taylor diagrams (Figure 5) are closest to the reference optimal point. This is consistent with their relatively high values of the correlation coefficients (> 0.7, but chlorophyll) and relatively low values of the unbiased RMSE in Table2.4. Furthermore, the above variables fall the closest to the reference (0,0) point in the Target diagrams (figure 15), consistently with the relatively low values of the percentage model bias and RMSE in Table 2.4. Finally, we observe that the reliability index value in Table 2.4 is close to one for all the above variables (but chlorophyll), indicating that the model predictions has the same order of magnitude of the data.

The model had some skill in hindcasting nutrients according to some of the metrics, despite the negative model efficiency. We note in particular the not negligible correlation values for silicate, phosphate and nitrate (see Table 2.4 and the Taylor diagrams).

However, the model estimates of nutrients had relatively high values of the percentage model bias (Table 2.4) driving their representative points outside the targets in Figure 6. In particular, the model had the tendency of overestimating the in situ data, as highlighted by the negative values of the percentage % bias in Table 2.4, and the location in the I and IV quadrants of the target points (Figure 6).

To this regard, the density plots in Figure 14 show that the model had the tendency to overestimate low values data. This is particularly evident for nitrate, for which the model produced a large range of estimates stacked to a constant measurement value of 0.1 mmol m^{-3}. This outcome may indicate also that detection limits and precision of measurements of this variable may have affected the benchmarking results for nutrients and nitrate in particular.

Also the relatively low values of the Pearson correlations of nutrients can be interpreted looking at the density plots in Figure 4. The highest densities of the match-ups fell approximately on a 1:1 line for nitrate, phosphate and silicate, suggesting, qualitatively, a linear relation among model and data. However, low densities are distributed towards high values of model output. This may have lowered the values of the Pearson correlation coefficient, since it is a parametric statistics which may be markedly affected by extreme match-up values.

Finally, we note the relatively low model bias for the pH estimates (Table 2.4). However, the values of the other skill metrics indicate a relatively poor model performance with respect to this variable.

1.7 Summary of hindcast experiment and lessons learnt

Overall the benchmarking exercise showed reasonable results for the main components of the modelled ecosystem given the different character of the in-situ measurements (instantaneous, punctual, episodic) and model outputs (bulk quantities for substantial water masses averaged over a given time period), however with some margins for improvements.

The correlations from the weakest (~0.2 for pH and ammonium) to the highest (0.9 for temperature) indicate that the general functioning of the ecosystem are captured adequately, while some other results of the statistics, e.g. the large bias for nitrogen, suggest focal points for future model analysis, like the individuation of "hot-spots" of poor model performance and the investigation of the dominant processes in these areas as candidates for model improvement.

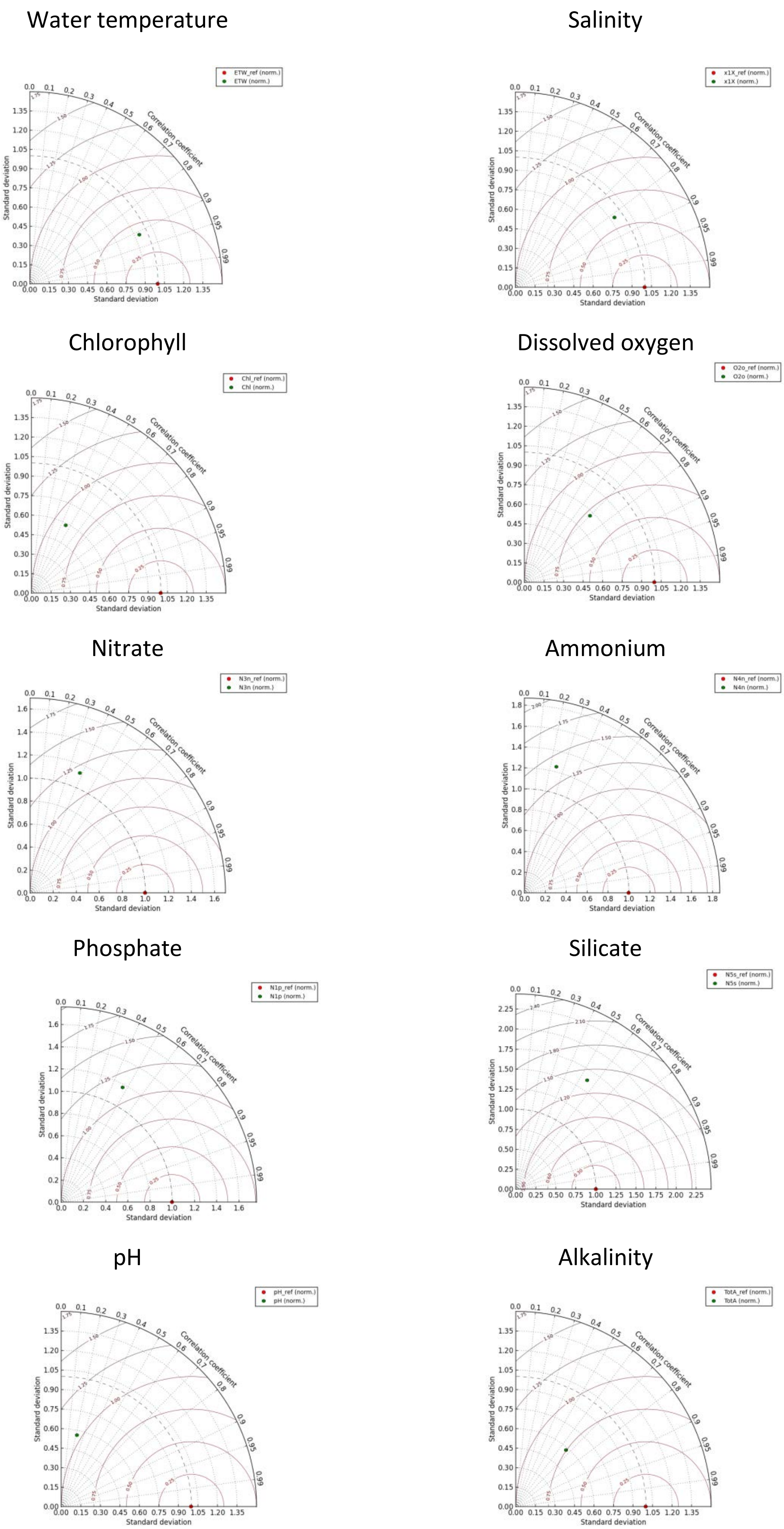


Figure 5. Taylor diagrams for the ten validation variables.

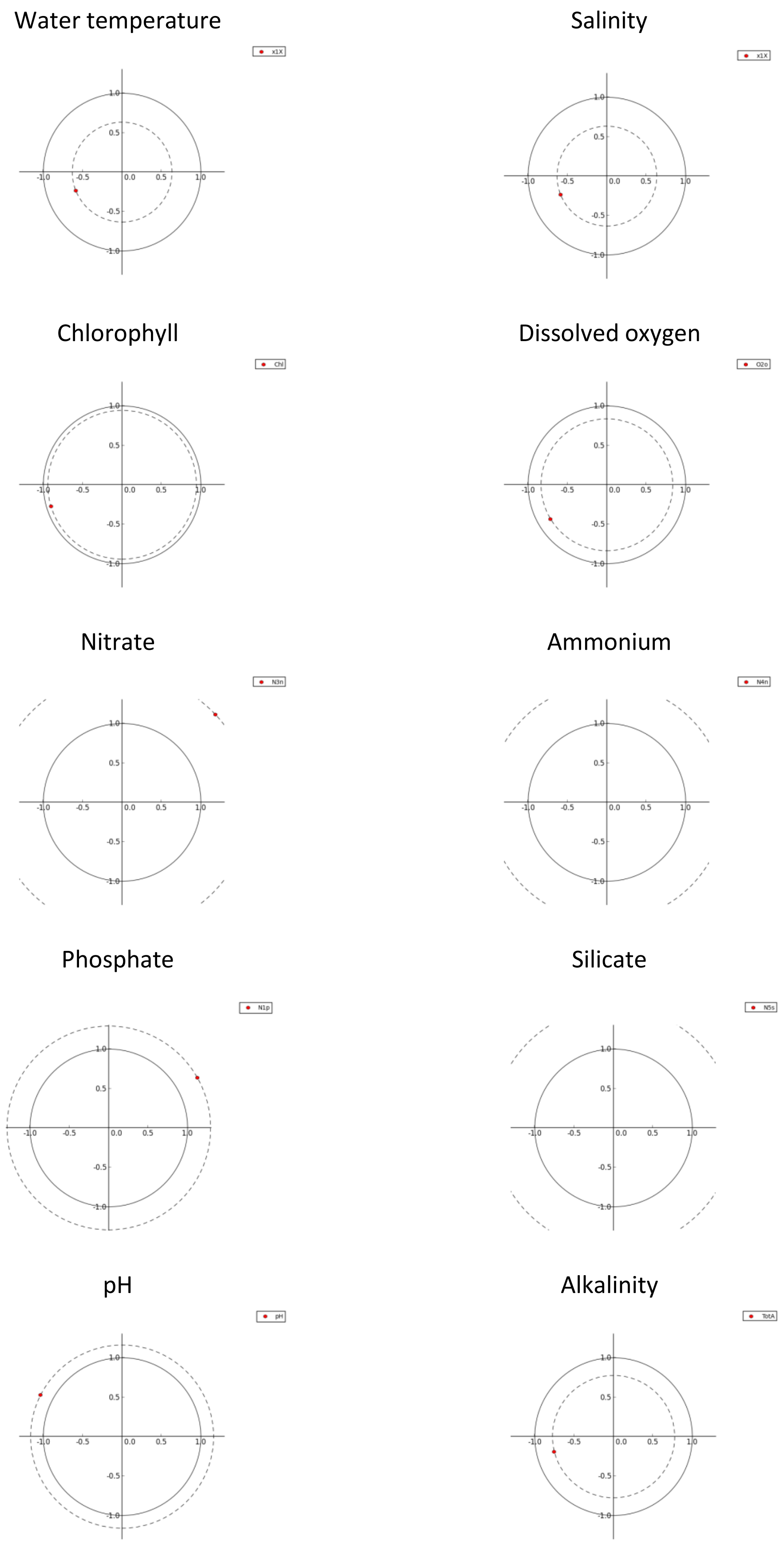


Figure 6. Target diagrams for the ten validation variables.

1.8 References

ICES, 2009. ICES Dataset on Ocean Hydrography. The International Council for the Exploration of the Sea, Copenhagen.

Allen, J. I.; Blackford, J. C.; Holt, J.; Proctor, R.; Ashworth, M. & Siddorn, J. A highly spatially resolved ecosystem model for the North West European Continental Shelf Sarsia, 2001, 86, 423-440

Artioli, Y.; Blackford, J. C.; Butenschön, M.; Holt, J. T.; Wakelin, S. L.; Thomas, H.; Borges, A. V. & Allen, J. I. The carbonate system in the North Sea: Sensitivity and model validation Journal of Marine Systems, 2012, 102-104, 1 – 13

de Mora, L.; Butenschön, M. & Allen, J. I. How should sparse marine in situ measurements be compared to a continuous model: an example Geoscientific Model Development, 2013, 6, 533-548

Holt, J.; Butenschön, M.; Wakelin, S. L.; Artioli, Y. & Allen, J. I. Oceanic controls on the primary production of the northwest European continental shelf: model experiments under recent past conditions and a potential future scenario Biogeosciences, 2012, 9, 97-117

Lenhart, H.-J.; Mills, D. K.; Baretta-Bekker, H.; van Leeuwen, S. M.; van der Molen, J.; Baretta, J. W.; Blaas, M.; Desmit, X.; Kühn, W.; Lacroix, G.; Los, H. J.; Ménesguen, A.; Neves, R.; Proctor, R.; Ruardij, P.; Skogen, M. D.; Vanhoutte-Brunier, A.; Villars, M. T. & Wakelin, S. L. Predicting the consequences of nutrient reduction on the eutrophication status of the North Sea Journal of Marine Systems, 2010, 81, 148 – 170

Pätsch, J. and Lenhart, H.-J.: Daily loads of nutrients, total al-akalinity, dissolved inorganic carbon and dissolved organic carbon of the European continental rivers for the years 1977–2002, Berichte aus dem Zentrum f˙ r Meeres- und Klimaforschung; Reihe B: Ozeanographie, 48, 159pp., 2004.

Siddorn, J. R.; Allen, J. I.; Blackford, J. C.; Gilbert, F. J.; Holt, J. T.; Holt, M. W.; Osborne, J. P.; Proctor, R. & Mills, D. K. Modelling the hydrodynamics and ecosystem of the North-West European continental shelf for operational oceanography Journal of Marine Systems, ELSEVIER SCIENCE BV, {2007}, {65}, {417-429}

2. Ecospace driven by POLCOMS-ERSEM

2.1 Description NE Atlantic

The Ecospace North Sea Model run was between 1991 and 2009. The first 8 years are treated as a burn in period to allow the Ecospace Model to stabilise. The Model consisted of 70 functional groups including detritus and producers and 5 multi-stanza double groups of adults and juveniles. The area covered was from 51 to 60 degrees latitude by 1/9 degree and -6 to 10 longitude by 1/6 degree. Information on biomass of Phytoplankton, Herbivorous and Omnivorous Zooplankton as well as Microzooplankton and Bacteria and temperature, depth salinity information and was fed from the POLCOMS-ERSEM model (see section) on a daily basis. The Ecospace model is a mass balance model that emphasises the trophic relationships within a food web. Of the 70 groups, the benchmarked groups are indicated below, to include a range of commercially important and environmentally sensitive groups. The benchmarking uses fixed fishing mortality (with benchmark data detrended in the case of Cod) so that lower trophic level and food web effects alone are being studied. There is only one model output for seabirds but it has been benchmarked with scaled data for two species.

Table 2.1 Summary of the benchmarking of ecospace in the North Sea. Model skill scores are colour coded to give an indication of model skill as follows; Good = green, Moderate = Black, Poor = Red; correlation r good > 0.75, poor < 0.20, ME good = > 0.5, Poor = < 0, Reliability index good 0.8-1.2, poor > 2 or –ve. PBias, good = < 20%, poor = > 100%.

Region: NE Atlantic
Model: Ecospace driven by POLCOMS-ERSEM
Met forcing: DMI regional hindcasat
Hindcast time period: 1990-2009
Contact: Jonathan.Beecham@Cefas.co.uk (HTL only)

		Hindcast					
Variable	**Unit**	**Model Efficiency**	**% Model Bias**	**Pearson Correlation Coefficient**	**Spearman Correlation Coefficient**	**Unbiased RMSE**	**Reliability Index**
TSB Adult Herring	Tonnes	0.202	0*	0.362	0.391	582000	0.106
Herring Recruit	No.	**-0.398**	0*	**-0.138**	0.072	24200000	**-0.0873**
Cod Adult TSB +	Tonnes	**-0.184**	0*	**-0.0723**	0.000	47800	**-1.07**
Cod Recruit +	Tonnes	**-0.0195**	0*	**0.0189**	0.182	68800	**-0.460**
Sandeel TSB	Tonnes	**-0.731**	0*	**-0.485**	-0.595	338000	**-0.119**
Plaice TSB	Tonnes	**-0.737**	0*	**-0.591**	-0.691	70000	**-0.449**
Guillimot Abundance Index	%	**-0.257**	0*	**0.0860**	0.173	18.4	**-3.15**
Black Footed Kittiwake A.I	%	**0.806**	0*	**0.895**	0.773	4.44	**0.797**

Notes - * - Bias is zero because means of observed are scaled to means of data. + - detrended data.

Table 2.2

Variable	Unit	Model Correlation Time (years)	Data Correlation Time (years)	Model Year 1 autocorrelation	Data Year 1 autocorrelation
TSB Adult Herring	Tonnes	0	2	-0.152	0.693
Herring Recruit	No.	0	2	-0.367	0.219
Cod Adult TSB	Tonnes	3	2	0.742	0.468
Cod Recruit	No.	3	0	0.765	-0.248
Sandeel TSB	Tonnes	1	0	0.328	-0.072
Plaice TSB	Tonnes	2	2	0.592	0.460
Guillimot Abundance Index	%	2	0	0.647	0.002
Black Footed Kittiwake A.I	%	2	3	0.647	0.458

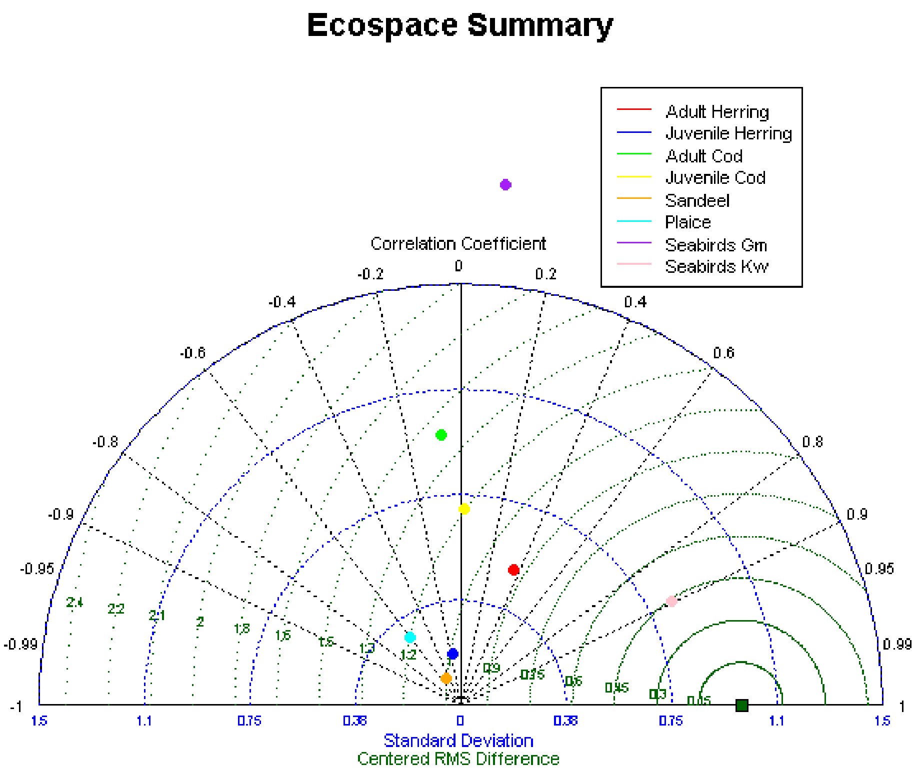


Figure 7. Taylor Diagram for the ecospace model of the North Sea.

2.2 Summary and lessons learnt

The results (Tables 2.2.1, 2.2.2, Figure 7) show modest positive or negative correlations for most datasets. This is to be expected as only some of the effects of lower trophic levels on the HTL N Sea model are included. In future model iterations we might include effects of timing of larval bloom on spawning and hence future recruitment and direct effects of temperature on feeding and survival of HTL species independent of effects via LTL species.

In the first instance it is important to ensure that the time response, as indicated by the autocorrelation is correct. Generally the response was acceptable for adult Cod, Kittiwakes and Plaice but modelled Herring was over responsive to short term signals in drivers and modelled Guillemot and juvenile Cod are under responsive. The former is probably spurious since this is a long –lived relatively slow growing species and the short term variation may reflect the noise in the data collection. However there is a serious problem in Ecospace and Ecosim in that multi-stanza models show strong correlation between stanza biomasses and do not show the correct differential sensitivity to the frequency of the input signal.

There is a tendency for the model to underestimate the RMSE of the data (apart form in adult cod and seabirds), indicating that the model is not capturing all the sources of variation. In particular the model does not predict any major source of population variation for Plaice (not surprising as there is no direct driving for HTL benthic species) and Sandeels, which is a species where we would expect considerable population variability given the frequency of the data. Specific mechanisms of variation in populations, such as match-mismatch, need to be examined for these species. The surprisingly high correlations for Kittiwakes are probably partly co-incidental in that both model and data happened to have a downward trend, but do indicate a good frequency response to input signal data.

Current developments in the Ecospace model will allow a wider range of influences between LTL drivers and HTL population dynamics to be modelled and should improve over the lifetime of this project and beyond. However it is encouraging that there seems to be a positive correlation between input data and HTL response for Herring, which is an important small pelagic species, which seems to be highly sensitive to short-term variation in food supply (Figure 8). We also note that there is quite a long settling time for the model – so that the predictions should be improved by starting the model run earlier.

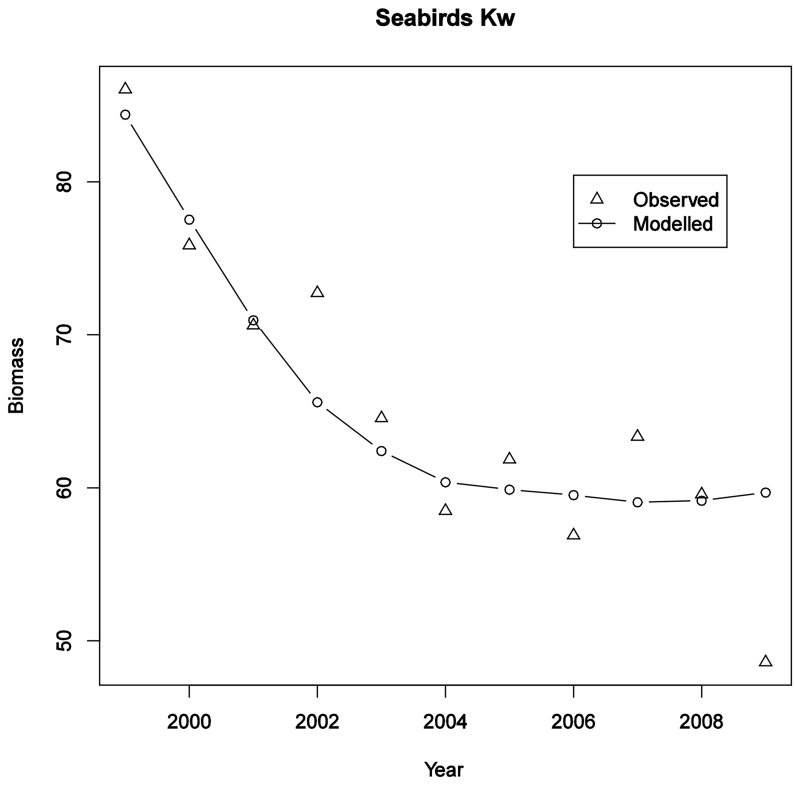


Figure 8. Comparison of modelled seabird biomass with observations.

2.3 Size-based modelling of higher trophic levels (Cefas, PML)

2.3.1 Description

The size-based model described in Blanchard et al (2008) was also applied to the POLCOMS-ERSEM hindcast from PML. This model incorporates the two main carbon pathways in the higher trophic levels, represented by pelagic predators (dependent on a size-structured food source) and detritivores (dependent on a non-size structured food source). Depth-integrated values for three phytoplankton functional groups and two zooplankton functional groups from ERSEM were used as a food source for pelagic predators, while ERSEM generated detritus (bottom layer) served as a food source for detritivores. The simple representation of size classes provides an effective way to describe higher trophic levels without the uncertainty related to species-specific parameters. Fishing pressure is included as is temperature effects on feeding. The size-based model has been adapted to run on a daily timescale, so that seasonal effects from the lower trophic levels feed through into the higher trophic levels. This allows the model to be applied for the Rapid Environmental Assessments, for which it will provide information on biomass, fish yield and the balance between the two carbon pathways.

2.3.2 Validation

Validation of the size-based model is hampered by a lack of validation data aggregated on the size-spectrum scale. Observations for pelagic predators were used from Jennings et al (2002), while detritivore observations came from Maxwell & Jennings (2006). These were compared with modelled results for the same area but covering the entire hindcast period. Table 2.3.1 contains the validation results, while Figure 22 shows the corresponding Taylor diagram.

Table 2.3.1

Region: **North East Atlantic**
Model: size-based model (Blanchard et al, 2008)
Met forcing: from POLCOMS-ERSEM hindcast
Hindcast time period: 1991-2009
Contact: Sonja van Leeuwen sonja.vanleeuwen@cefas.co.uk

		Hindcast				
Variable	**Unit**	**Model Efficiency**	**% Model Bias**	**Pearson Correlation Coefficient**	**RMSE**	**Unbiased RMSE**
Pelagic predators	log 10 (g wet weight / m^2)	0.73	-30.67	0.97	0.74	0.55
Detritivores	log10 (g wet weight / m^2)	0.41	-71.16	0.87	0.39	0.35

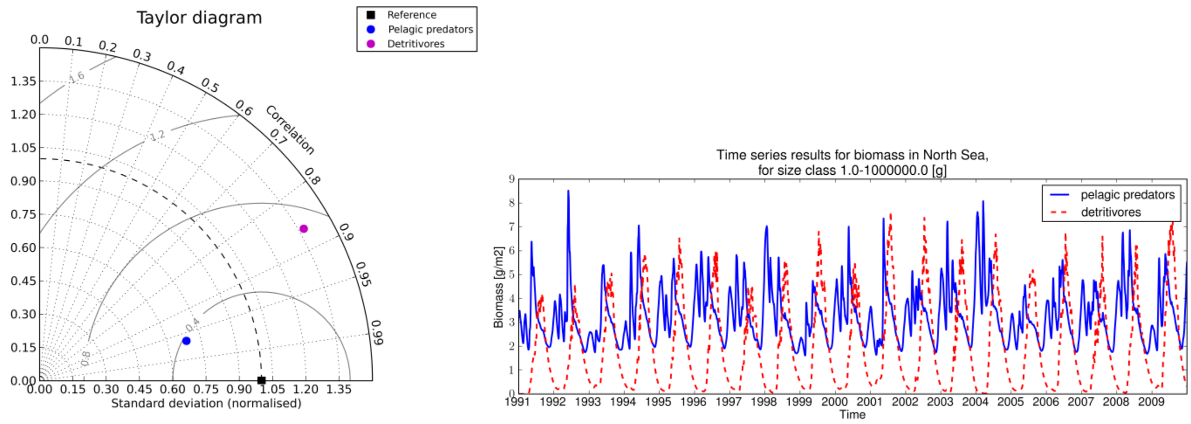


Figure 9. Taylor diagram of the size-based model results for log10 biomass (left) and biomass results for the North Sea, integrated over the 1g-1000kg size range (right).

2.3.3 Summary and lessons learnt

In all, model skill is better for pelagic predators (which span the entire modelled size-spectrum) than for detritivores (which occupy a part of the size-spectrum). Although correlation is high (reflecting that the marine ecosystem is a size-structured system), percentage bias can be large. This is due to fishing pressure, with the observations representing a heavily fished system. The model applies a standard fishing mortality, but has not been tuned to represent any area-specific fishing pressure. This is also visible in the resulting size-spectrums shown in Figure 10: the observations show less pelagic predators of commercial size and more small pelagic fishes. Detritivore biomass is high in the shallow southern North Sea, with pelagic predators dominant in the deeper northern North Sea (Figure 9). The large intra-annual variation in biomass is due to the smaller size-classes.

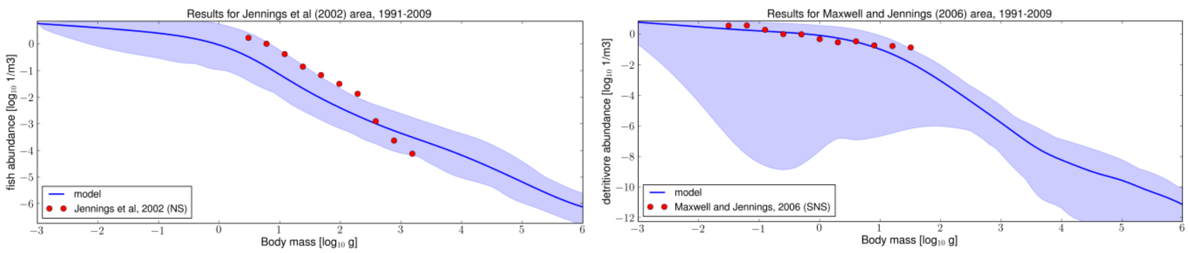


Figure 10. The modelled size-spectrum for pelagic predators (right) and detritivores (left), with the corresponding observations. The blue area denotes the model envelope.

2.4 References

Jennings, S., Warr, K.J., Mackinson, S. (2002) Use of size-based production and stable isotope analyses to predict trophic transfer efficiencies and predator-prey body mass ratios in food webs. Marine Ecology Progress Series, 240, pp. 11-20.

Maxwell, T.A.D. & Jennings, S. (2006) Predicting abundance-body size relationships in functional and taxonomic subsets of food webs. *Oecologia*, 150 (2), 282–290, DOI 10.1007/s00442-006-0520-2

Blanchard, J.L., Jennings, S., Law, R., Castle, M.D., McCloghrie, P., Rochet, M-J., Benoit, E., (2009), How does abundance scale with body size in coupled size-structured food webs? Journal of Animal Ecology, 78, pp. 270—280, doi: 10.1111/j.1365-2656.2008.01466.x