Validiting the COMFORT simulation

Robert Wilson | Plymouth Marine Laboratory

Apr 22, 2025

CONTENTS

Simulations were compared using the Python package **ecoval**.

You can find installation instructions for **ecoval** here.

A website with more information on **ecoval** can be found here.

ecoval is developed by Robert Wilson at Plymouth Marine Laboratory, who can be contacted at rwi@pml.ac.uk.

CHAPTER

ONE

MODEL RUN INFORMATION

For the purpose of our study we use the 3-D coupled hydrodynamic-biogeochemical modeling system NEMO-FABM-ERSEM configured on the Atlantic Meridional Margin (AMM7) domain, which is based on an update of Edwards et al. (2012). The geographical domain extends from 20°W to 13°E and 40°N to 65°N (Figure 1), and has a horizontal resolution of 1/15° in latitudinal and 1/9° in longitudinal direction, corresponding to ~7 km. Vertically, the model resolution was improved to 50 σ-z layers, from the original 32, and has a minimum bathymetry of 10 m. In areas of steep bathymetry change the total number of layers is decreased to reduce spurious vertical transport.

The model simulation began on **1st of January 1981** and ended on **31st of December 2017**.

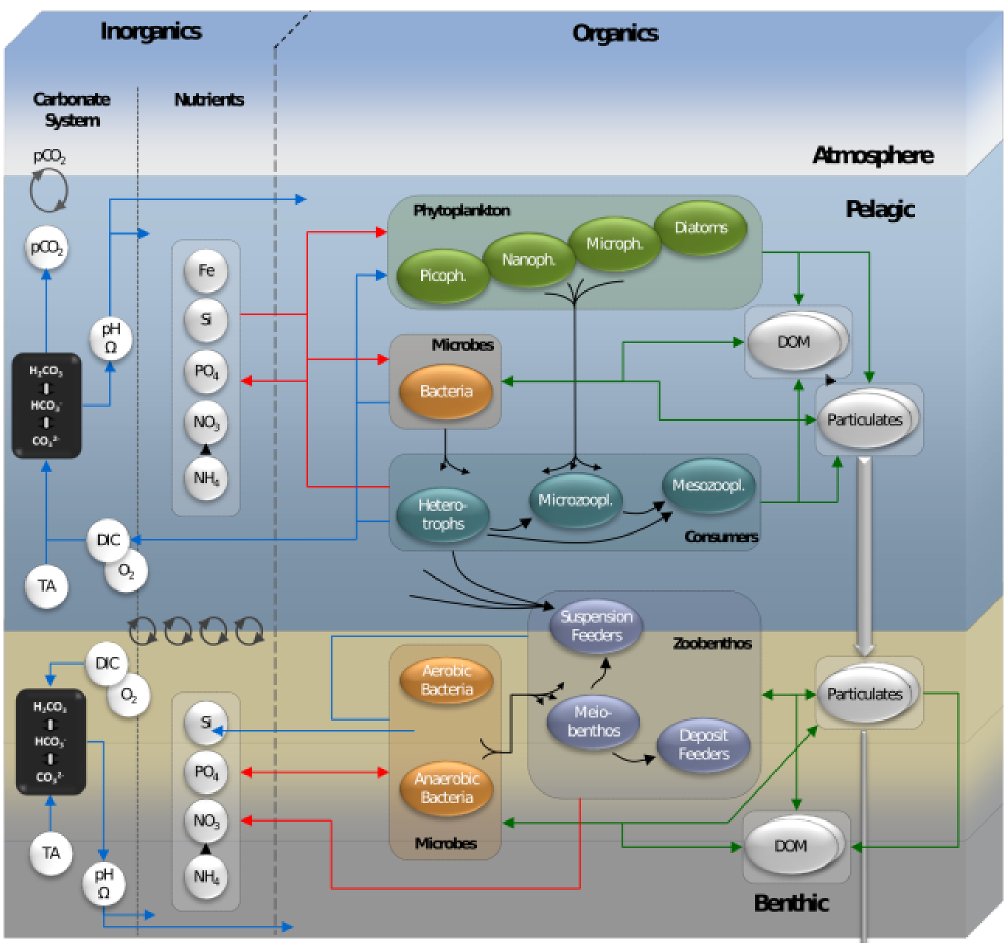


Figure 1.1: ERSEM schematic showing how model components interact with or influence each other. Blue connectors represent inorganic carbon fluxes, red represents nutrient fluxes, yellow represents oxygen, black represents predator–prey

interactions, and green represents fluxes of non-living organics. Dashed arrows indicate the influence of carbonate system variables.

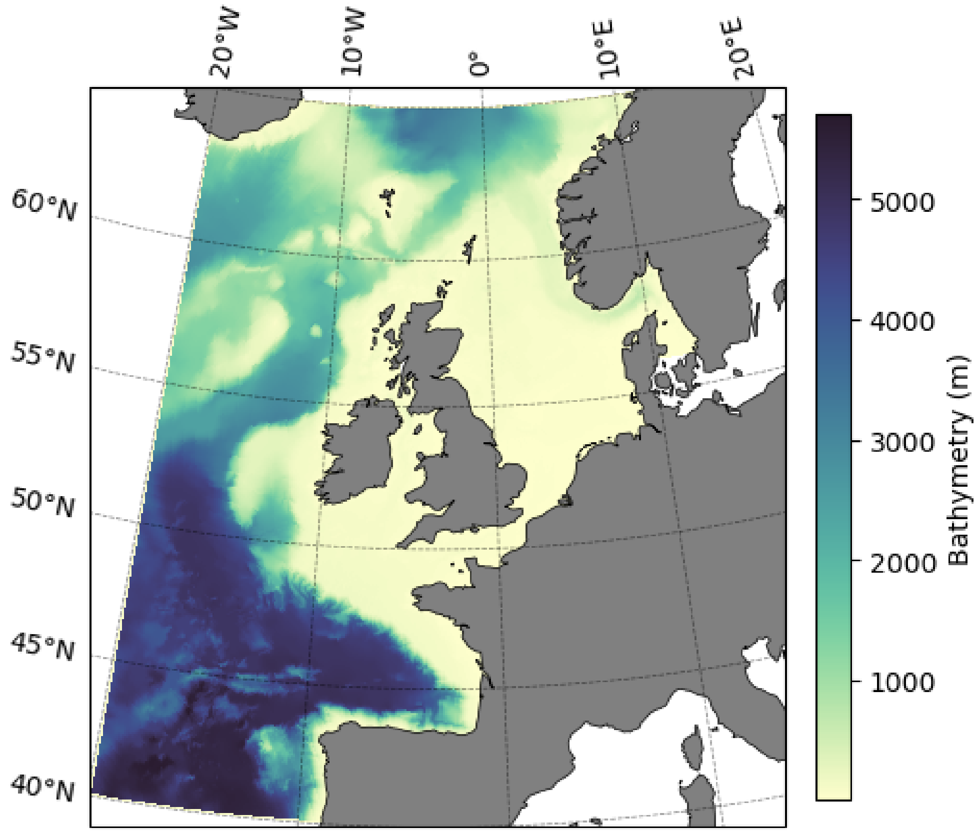


Figure 1.2: Bathymetry of the model domain. The model domain extends from -20° to 13° longitude and 40° to 65° latitude.

CHAPTER

TWO

METHODS

2.1 Definitions

Assume that we have a model m and observations o of the same size N.

Root mean square deviation (RMSD)

$$\text{RMSD}(m, o) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(m_i - o_i)^2}$$

Bias

$$\text{Bias}(m, o) = \frac{1}{N}\sum_{i=1}^{N}(m_i - o_i)$$

Pearson correlation coefficient

$$\text{Corr}(m, o) = \frac{\sum_{i=1}^{N}(m_i - \bar{m})(o_i - \bar{o})}{\sqrt{\sum_{i=1}^{N}(m_i - \bar{m})^2 \sum_{i=1}^{N}(o_i - \bar{o})^2}}$$

where $\bar{m}$ and $\bar{o}$ are the mean values of m and o respectively.

2.2 Software used

Models were matched up with observational data using the Python package nctoolkit. Plots were created using the Python package nctoolkit, and the R package ggplot2. Data analysis was performed using the Python packages pandas, nctoolkit, and statsmodels.

FULL DOMAIN SUMMARY STATISTICS OF MODEL PERFORMANCE

3.1 Taylor diagrams for the sea surface

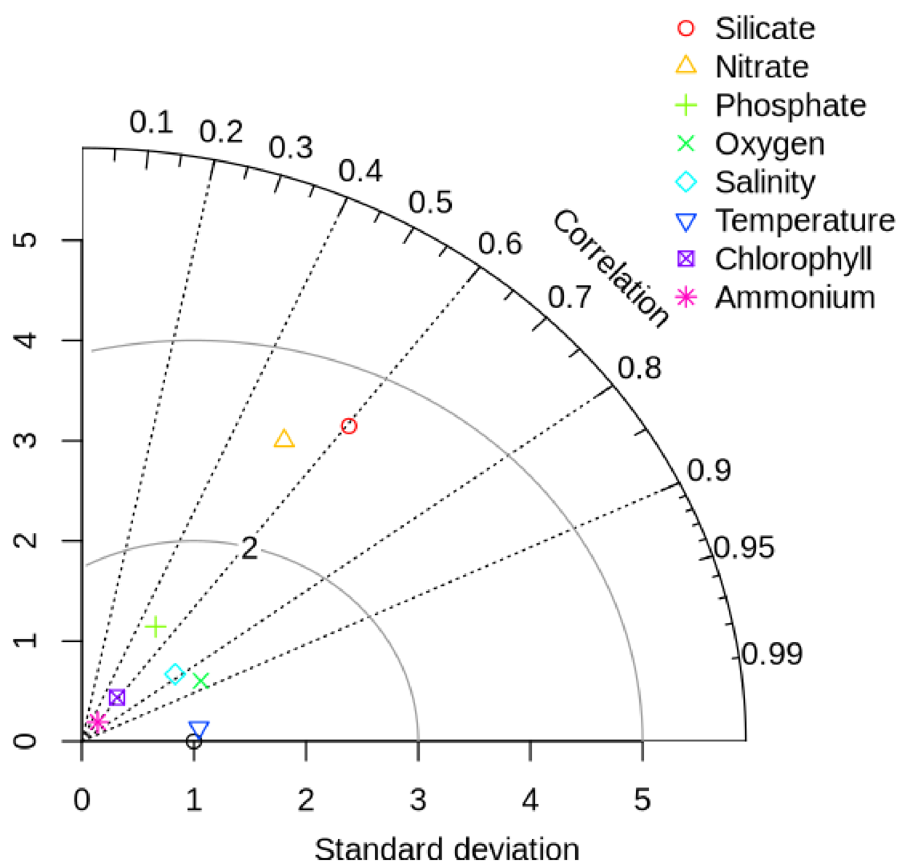


Figure 3.1: Taylor diagram for **sea surface** annual mean of Ammonium, Chlorophyll, Nitrate, Oxygen, Phosphate, Salinity, Silicate, Temperature. This diagram compares climatological annual averages of the model and observations across the model's spatial domain. Standard devaiation is normalized by the standard deviation of the observations, and a standard deviation below 1 indicates that the model is less variable than the observations. Note: This figure summarizes the overall ability of the model to reproduce climatological **spatial patterns**, and it does not represent temporal performance.

3.2 Model biases based on gridded sea surface data

Variable	Model mean	Observed mean	Model bias	Percentage bias
Ammonium	0.35	1.16	-0.80	-69.34
Chlorophyll	0.95	1.00	-0.04	-4.51
Nitrate	7.76	5.75	2.00	34.86
Oxygen	288.65	284.01	4.64	1.64
Phosphate	0.43	0.46	-0.04	-8.07
Salinity	34.82	34.80	0.02	0.07
Silicate	4.62	3.30	1.32	39.98
Temperature	11.94	12.23	-0.30	N/A

Table 3.1: Bias of model compared with **sea surface** observations. The bias is calculated as the modelled spatial mean minus the observational spatial mean. The percentage bias is calculated as the model bias divided by the observational spatial mean.

3.3 Spatial performance of the model at the sea surface

Variable	Spatial correlation between model and observations
Ammonium	0.61
Chlorophyll	0.59
Nitrate	0.54
Oxygen	0.87
Phosphate	0.48
Salinity	0.78
Silicate	0.61
Temperature	0.99

Table 3.2: Pearson correlation coefficient between model and observations at the **sea surface** for annual mean of ammonium, chlorophyll, nitrate, oxygen, phosphate, salinity, silicate, and temperature. This table compares climatological annual averages of the model and observations across the model's spatial domain. Standard devaiation is normalized by the standard deviation of the observations, and a standard deviation below 1 indicates that the model is less variable than the observations.

3.4 Temporal performance of the model at the sea surface

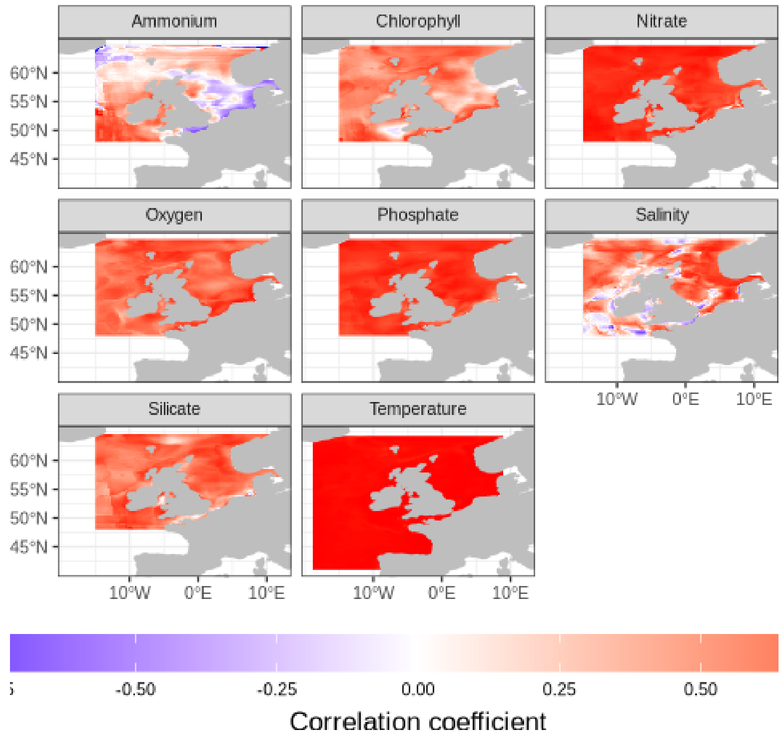


Figure 3.2: Spatial correlation (Pearson correlation coefficient) between model and observations for annual mean of ammonium, chlorophyll, nitrate, oxygen, phosphate, salinity, silicate, and temperature. This figure compares climatological monthly averages of the model and observations across the model's spatial domain.

The overall ability of the model reproduce the seasonality of each variable was estimated by calculating the spatial mean of the Pearson correlation coefficient between the model and the observations. The spatial mean was calculated by averaging the correlation coefficient of each grid cell.

Variable	r
Ammonium	0.23
Chlorophyll	0.58
Nitrate	0.92
Oxygen	0.76
Phosphate	0.91
Salinity	0.44
Silicate	0.69
Temperature	1.00

Table 3.3: Spatial average of the temporal correlation (Pearson correlation coefficient) between model and observations for annual mean of ammonium, chlorophyll, nitrate, oxygen, phosphate, salinity, silicate, and temperature. The correlation is calculated for each grid cell individually using monthly climatological averages. The spatial average is then calculated for each variable.

3.5 Perfomance of model across depths

Root mean square deviation (RMSD), bias and correlation between model and observations were calculated for each variable at different depths. The RMSD is calculated as the square root of the mean of the squared differences between the model and observations. The bias is calculated as the modelled spatial mean minus the observational spatial mean. The correlation is calculated for each variable at different depths.

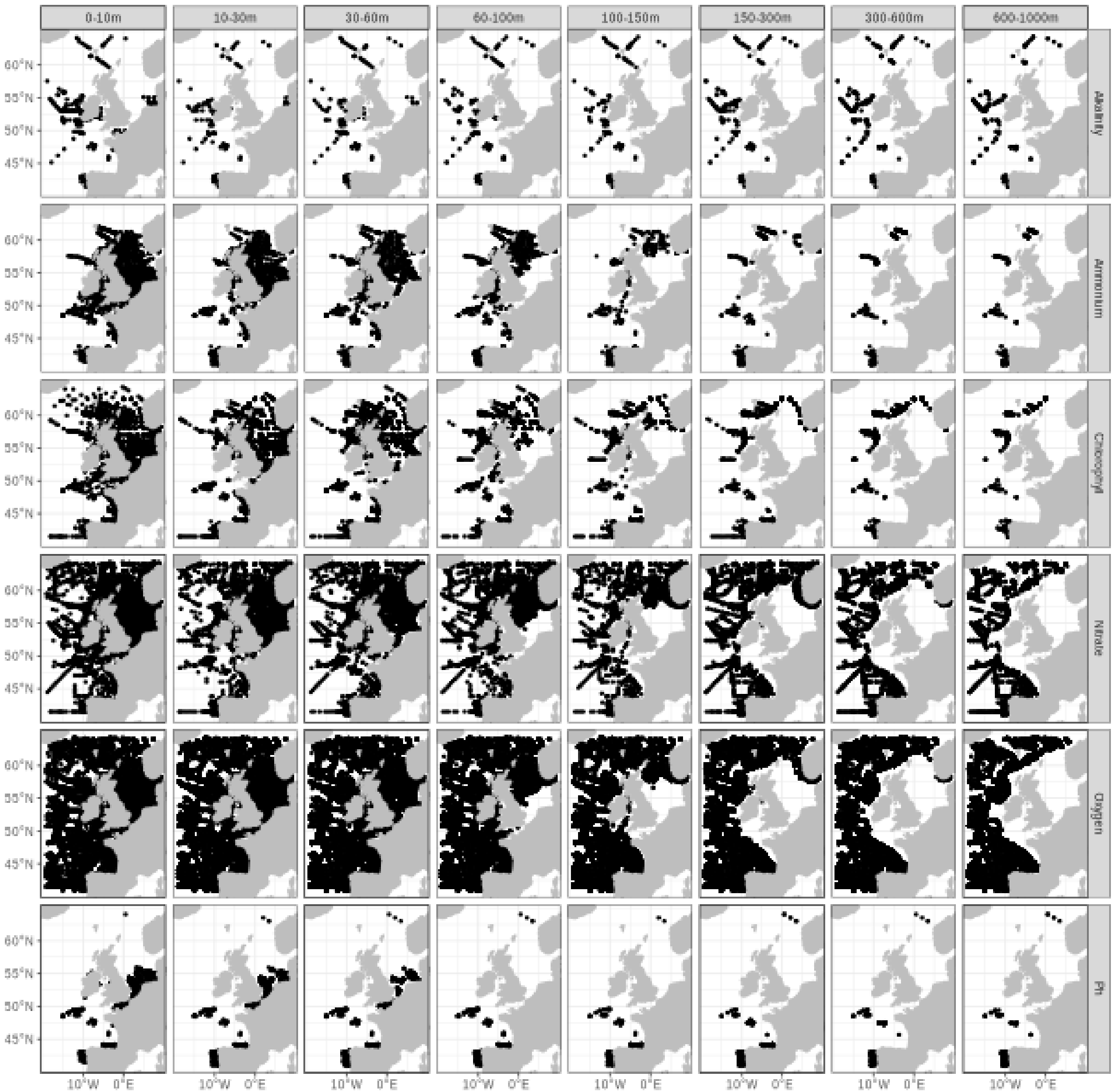


Figure 3.3: Map of the locations of matchups at each depth range.

Variable	Unit	0-10m	10-30m	30-60m	60-100m	100-150m	150-300m	300-600m	600-1000m	>1000m	0-150m
Alkalinity	mmolm^{-3}	66.63	41.41	32.31	29.72	36.67	38.65	43.03	37.14	46.71	46.27
Ammonium	mmol Nm^{-3}	1.42	1.17	1.22	1.15	0.64	0.77	0.90	0.92	0.69	1.28
Chlorophyll	mg m^{-3}	2.78	2.31	1.34	0.55	0.42	0.35	0.41	0.47	0.48	1.92
Nitrate	mmol Nm^{-3}	9.93	8.64	4.83	3.15	2.52	2.06	2.02	2.19	3.11	7.65
Oxygen	mmol O_2m^{-3}	31.49	33.63	30.32	17.31	12.79	11.29	13.29	25.04	17.24	25.88
Ph	-	0.18	0.20	0.21	0.22	0.21	0.22	0.09	0.08	0.04	0.20

Table 3.4: Root mean square deviation (RMSD) of model compared with observations at different depths. The RMSD is calculated as the square root of the mean of the squared differences between the model and observations.

Variable	Unit	0-10m	10-30m	30-60m	60-100m	100-150m	150-300m	300-600m	600-1000m	>1000m	0-150m
Alkalinity	mmolm^{-3}	12.02	23.82	18.16	16.99	23.98	25.34	30.54	16.08	26.15	17.73
Ammonium	mmol Nm^{-3}	-0.47	-0.41	-0.56	-0.50	-0.19	-0.43	-0.59	-0.61	-0.36	-0.47
Chlorophyll	mg m^{-3}	-0.07	0.15	0.28	-0.09	-0.18	-0.19	-0.21	-0.26	-0.21	0.04
Nitrate	mmol Nm^{-3}	4.51	4.35	3.19	2.15	1.54	1.17	0.70	-0.40	-1.46	3.63
Oxygen	mmol O_2m^{-3}	14.77	18.30	14.71	4.73	1.51	-0.05	-0.25	16.59	-5.64	10.41
Ph	-	0.06	0.07	-0.00	-0.04	-0.02	0.14	0.01	-0.00	0.00	0.01

Table 3.5: Bias of model compared with observations at different depths.

Variable	0-10m	10-30m	30-60m	60-100m	100-150m	150-300m	300-600m	600-1000m	>1000m	0-150m
Alkalinity	-0.30	0.22	0.60	0.73	0.79	0.67	0.68	0.76	0.66	0.10
Ammonium	0.21	0.27	0.39	0.21	0.33	-0.08	0.08	0.22	0.29	0.25
Chlorophyll	0.26	0.19	0.07	0.02	-0.15	-0.08	-0.03	-0.09	-0.05	0.30
Nitrate	0.60	0.53	0.50	0.62	0.61	0.56	0.32	0.46	0.68	0.57
Oxygen	0.57	0.50	0.38	0.60	0.71	0.79	0.82	0.77	0.65	0.55
Ph	0.25	0.32	0.24	0.29	0.18	0.47	0.10	-0.33	-0.07	0.21

Table 3.6: Pearson correlation coefficient between model and observations at different depths. The correlation is calculated for each variable at different depths.

CHAPTER

FOUR

VALIDATION OF ALKALINITY USING POINT OBSERVATIONS

4.1 Performance of model sea surface alkalinity

Values from the **top 5m** of the water column were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 223 values extracted from the observational database. The map below shows the locations of the matched up data for alkalinity.

The following model output was used to compare with observational values: **O3_TA**.

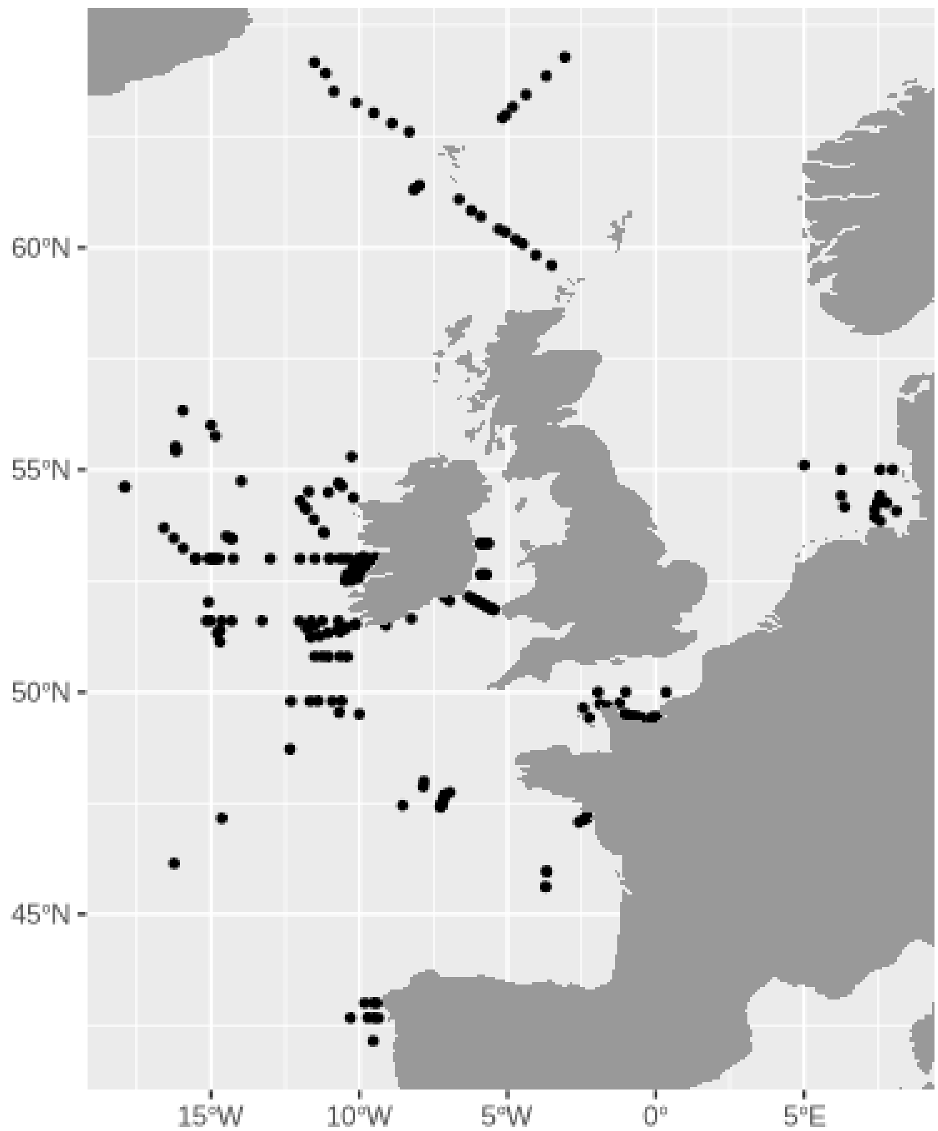


Figure 4.1: Locations of matchups between simulated and observed alkalinity in the top 5m of the water column.

The number of observations in each month ranged from 1 in December to 54 in February. Figure 4.2 below shows the distribution of observations in each month.

Figure 4.2 below shows the bias between the model and observational data for alkalinity. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

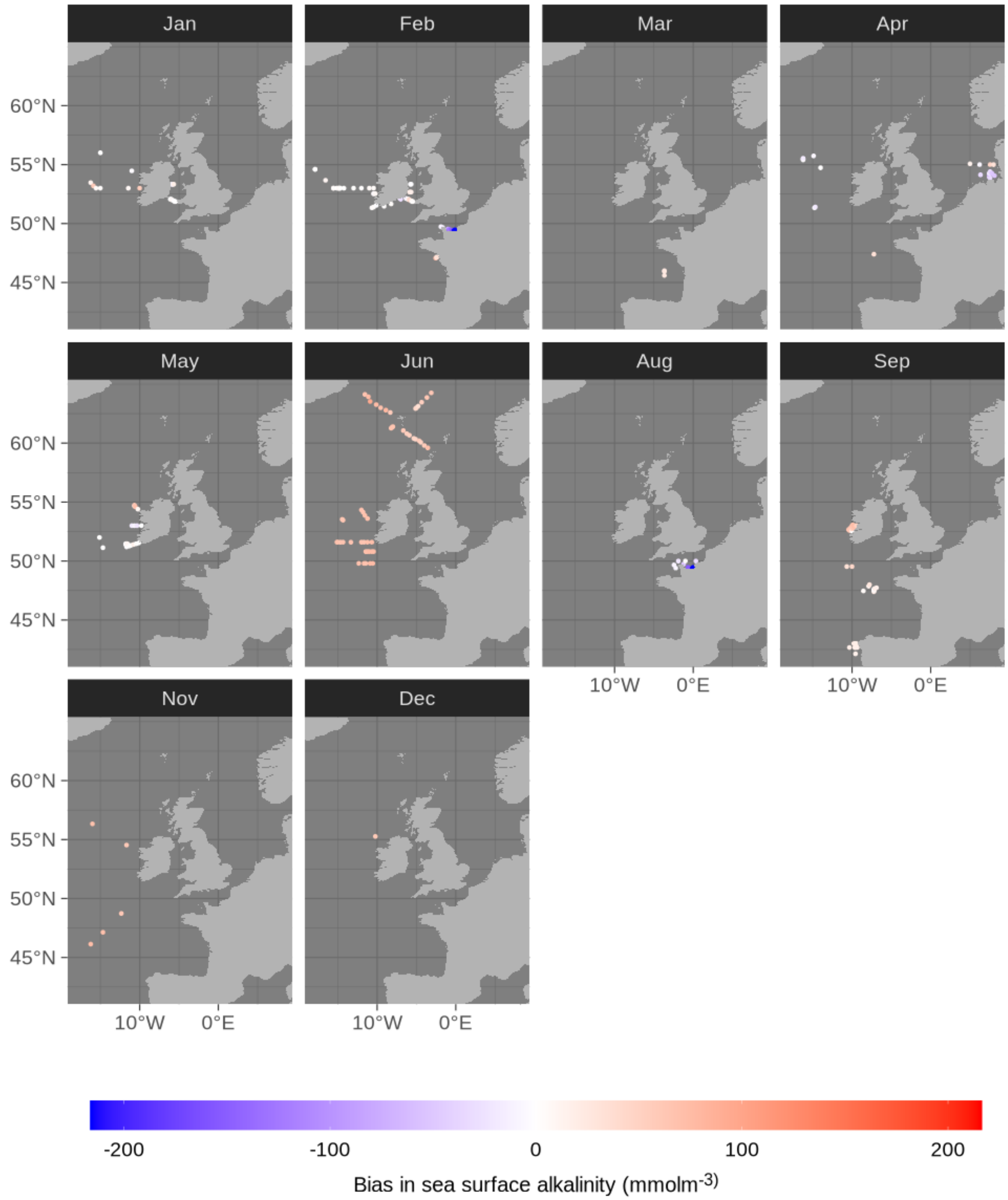


Figure 4.2: Bias in sea surface alkalinity. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 4.3 and 4.4 show the distribution of sea surface alkalinity observations in the model and observational datasets. This is shown for each month of the year (Figure 4.3) and for the entire year (Figure 4.4).

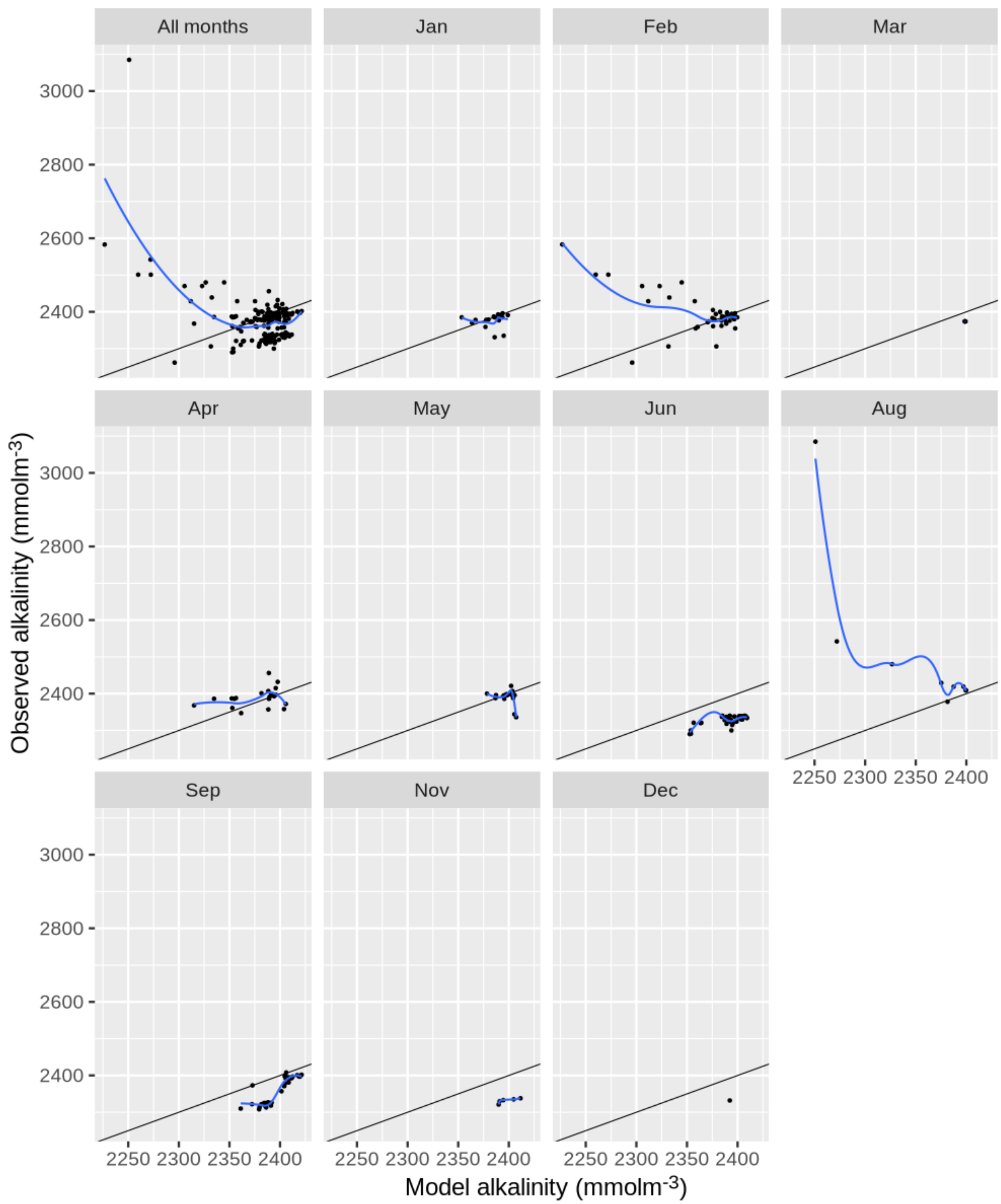


Figure 4.3: Simulated versus observed alkalinity in the top 5m of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

4.2 Summary statistics for sea surface alkalinity

The overall ability of the model to predict the observed alkalinity was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	12.78	83.27	-0.41	223
Jan	6.54	21.71	0.07	18
Feb	-22.54	80.62	-0.32	54
Mar	24.84	24.84	N/A	4
Apr	-3.11	48.93	0.05	22
May	6.91	22.78	-0.25	20
Jun	64.79	65.55	0.74	54
Aug	-153.35	297.54	-0.81	9
Sep	35.16	42.91	0.87	36
Nov	66.96	67.14	0.82	5
Dec	60.30	60.30	N/A	1

Table 4.1: Average bias (mmolm^{-3}) and root-mean square deviation (mmolm^{-3}) for the model's sea surface alkalinity for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed alkalinity was performed. The modelled alkalinity was used as the independent variable and the observed alkalinity was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	-1.10	4992.97	0.23	0.00
Jan	0.10	2130.48	0.00	0.79
Feb	-0.75	4173.82	0.34	0.00
Mar	-0.00	2374.00	-inf	1.00
Apr	0.33	1593.38	0.10	0.17
May	-0.65	3943.89	0.06	0.29
Jun	0.58	939.53	0.54	0.00
Aug	-3.12	9862.43	0.66	0.01
Sep	2.07	-2595.27	0.75	0.00
Nov	0.57	968.85	0.68	0.09

Table 4.2: Linear regression analysis of modelled and observed alkalinity. The modelled alkalinity was used as the independent variable and the observed alkalinity was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

4.3 Performance of model near-bottom alkalinity

Table 4.3: Average bias (mmolm^{-3}) and root-mean square deviation (mmolm^{-3}) for the model's near-bottom alkalinity for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

4.4 Depth-resolved comparisons of modelled and observed alkalinity

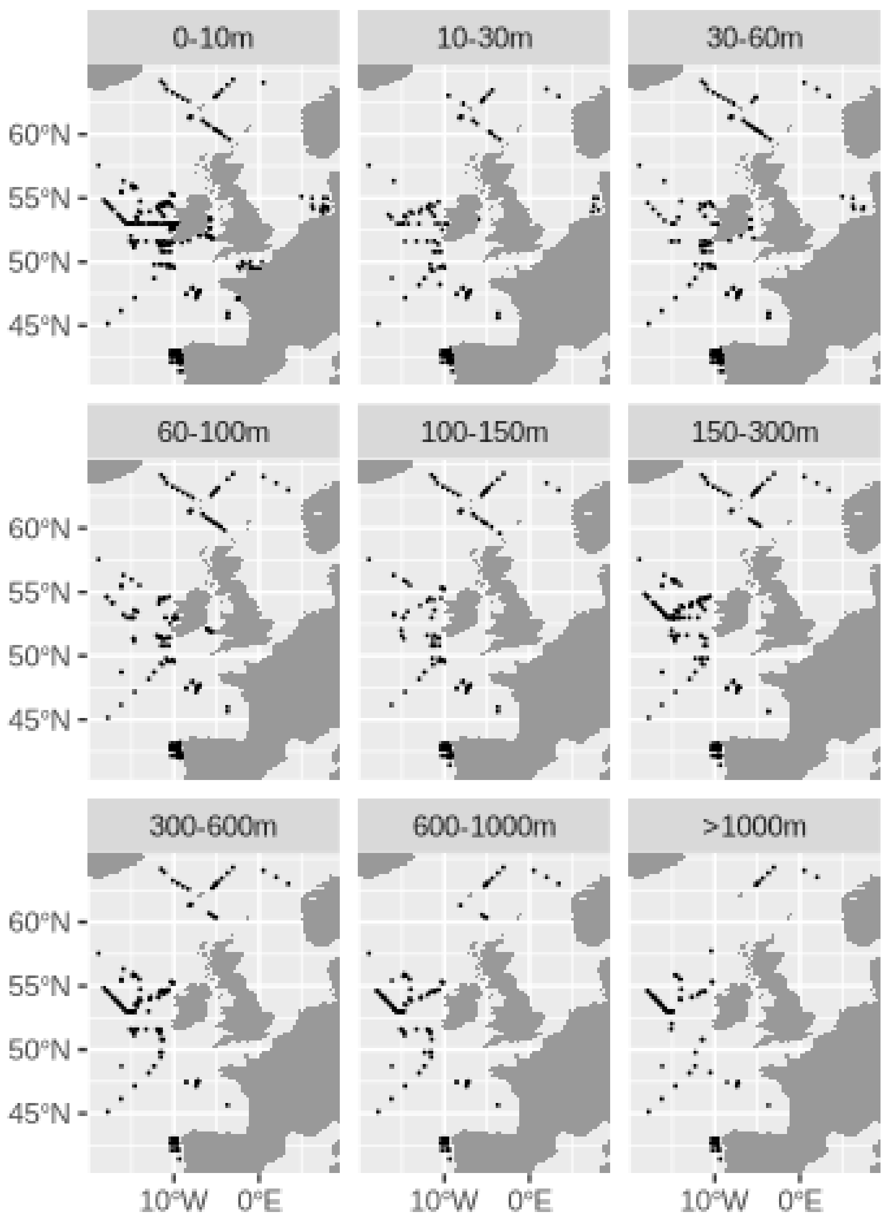


Figure 4.4: The geographic distribution of matchups between the model and observational alkalinity. The data has been binned into depth ranges. The depth ranges are 0-10m, 10-30m, 30-60m, 60-100m, 100-150m, 150-300m, 300-600m,

600-1000m, and >1000m. The number of observations in each depth range is shown in the tables below.

Depth	Bias	*r*	RMSD	Number of observations
0-150m	17.73	0.10	46.27	1,229
0-10m	12.02	-0.30	66.63	360
10-30m	23.82	0.22	41.41	172
30-60m	18.16	0.60	32.31	263
60-100m	16.99	0.73	29.72	260
100-150m	23.98	0.79	36.67	174
150-300m	25.34	0.67	38.65	236
300-600m	30.54	0.68	43.03	290
600-1000m	16.08	0.76	37.14	270
>1000m	26.15	0.66	46.71	296

Table 4.4: Average bias (mmolm^{-3}), root-mean square difference (RMSD) and correlation coefficient of modelled and observed alkalinity for different depth ranges. The bias is calculated as model-observation. The RMSD is the square root of the mean squared difference. The correlation coefficient is the Pearson correlation coefficient between the model and observed values.

4.5 Data Sources for validation of alkalinity

ICES Data Portal, Dataset on Ocean HydroChemistry, Extracted March 3, 2023. ICES, Copenhagen

CHAPTER

FIVE

VALIDATION OF AMMONIUM USING POINT OBSERVATIONS

5.1 Performance of model sea surface ammonium

Values from the **top 5m** of the water column were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 11,475 values extracted from the observational database. The map below shows the locations of the matched up data for ammonium concentration.

The following model output was used to compare with observational values: **N4_n**.

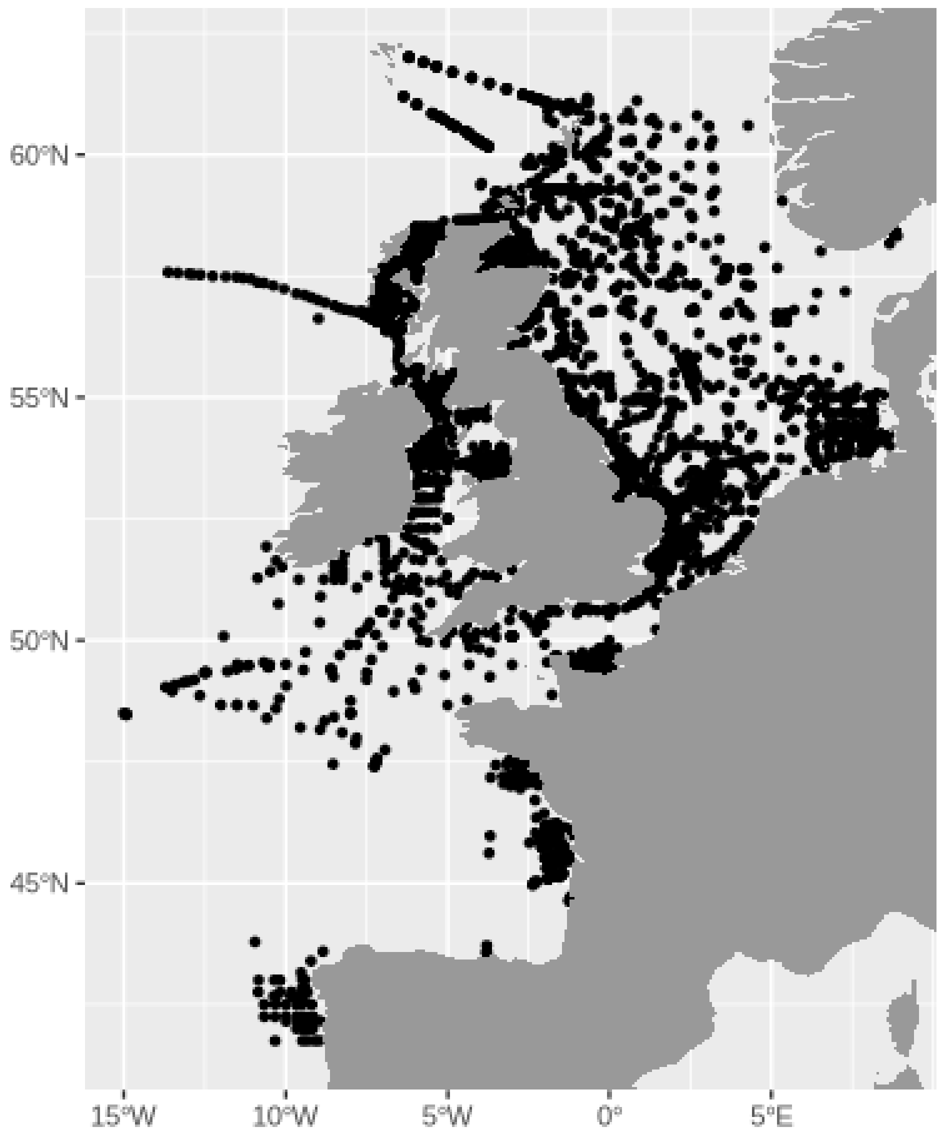


Figure 5.1: Locations of matchups between simulated and observed ammonium concentration in the top 5m of the water column.

The number of observations in each month ranged from 611 in November to 1,422 in January. Figure 5.2 below shows the distribution of observations in each month.

Figure 5.2 below shows the bias between the model and observational data for ammonium concentration. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

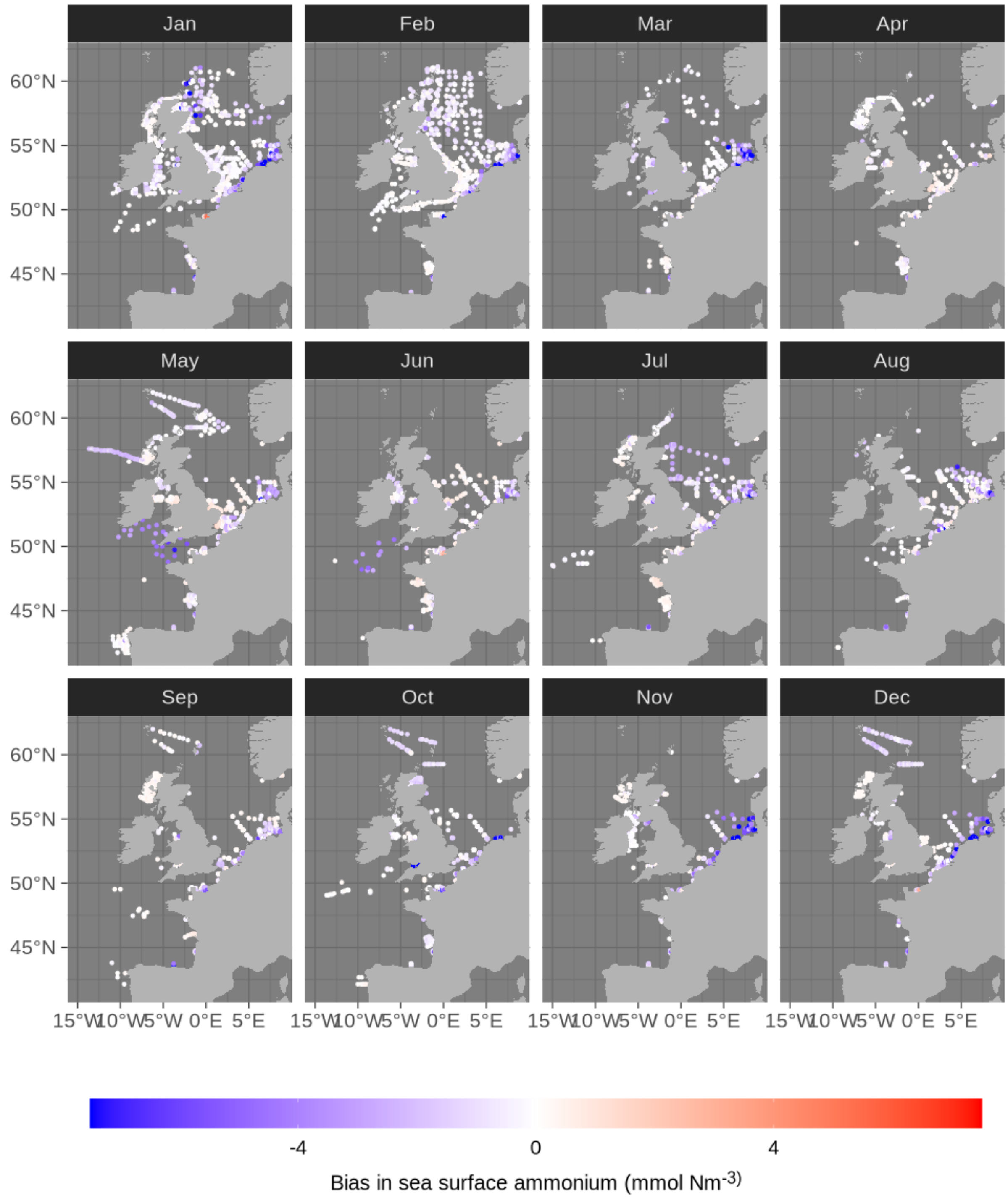


Figure 5.2: Bias in sea surface ammonium concentration. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 5.3 and 5.4 show the distribution of sea surface ammonium concentration observations in the model and observational datasets. This is shown for each month of the year (Figure 5.3) and for the entire year (Figure 5.4).

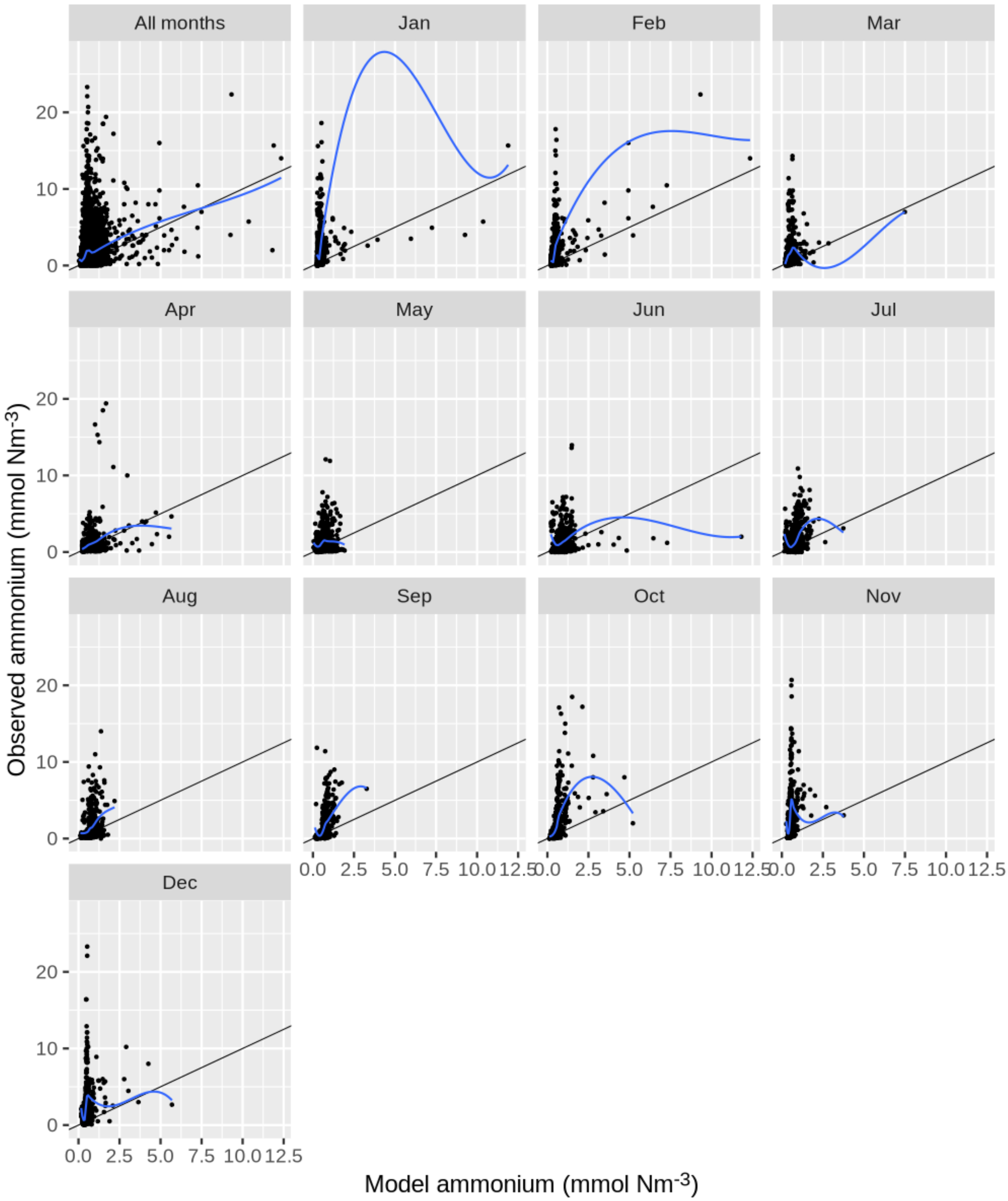


Figure 5.3: Simulated versus observed ammonium concentration in the top 5m of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

5.2 Summary statistics for sea surface ammonium

The overall ability of the model to predict the observed ammonium concentration was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	-0.85	2.18	0.22	11,475
Jan	-0.97	2.20	0.25	1,422
Feb	-0.81	2.02	0.49	1,287
Mar	-1.06	2.28	0.17	740
Apr	-0.35	1.82	0.24	1,015
May	-0.46	1.64	0.13	1,377
Jun	-0.42	1.79	0.09	1,030
Jul	-0.57	1.56	0.32	1,077
Aug	-0.63	1.81	0.30	849
Sep	-0.78	1.90	0.40	710
Oct	-1.36	2.73	0.38	704
Nov	-2.06	3.63	0.25	611
Dec	-1.83	3.23	0.22	653

Table 5.1: Average bias (mmol Nm^{-3}) and root-mean square deviation (mmol Nm^{-3}) for the model's sea surface ammonium concentration for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed ammonium concentration was performed. The modelled ammonium concentration was used as the independent variable and the observed ammonium concentration was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	1.03	0.88	0.07	0.00
Jan	0.86	0.99	0.07	0.00
Feb	1.60	0.50	0.27	0.00
Mar	0.89	1.09	0.03	0.00
Apr	0.84	0.37	0.09	0.00
May	0.73	0.69	0.03	0.00
Jun	0.19	1.13	0.01	0.03
Jul	1.72	0.12	0.13	0.00
Aug	2.03	-0.09	0.11	0.00
Sep	3.27	-0.76	0.21	0.00
Oct	2.43	0.58	0.17	0.00
Nov	2.90	1.21	0.06	0.00
Dec	1.43	1.57	0.05	0.00

Table 5.2: Linear regression analysis of modelled and observed ammonium. The modelled ammonium concentration was used as the independent variable and the observed ammonium concentration was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

5.3 Performance of model near-bottom ammonium

Near-bottom values of ammonium were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The near-bottom was defined as observations **within 2m of the seabed**. This was interpolated to the observational grid using the $GEBCO_2$ bathymetry dataset. Model values were interpolated to the observational dataset's longitudes and latitudes using 3D interpolation. **Note:** this analysis has been restricted to observations on the shelf region. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 919 values extracted from the observational database. The map below shows the locations of the matched up data for ammonium concentration.

The following model output was used to compare with observational values: **N4_n**.

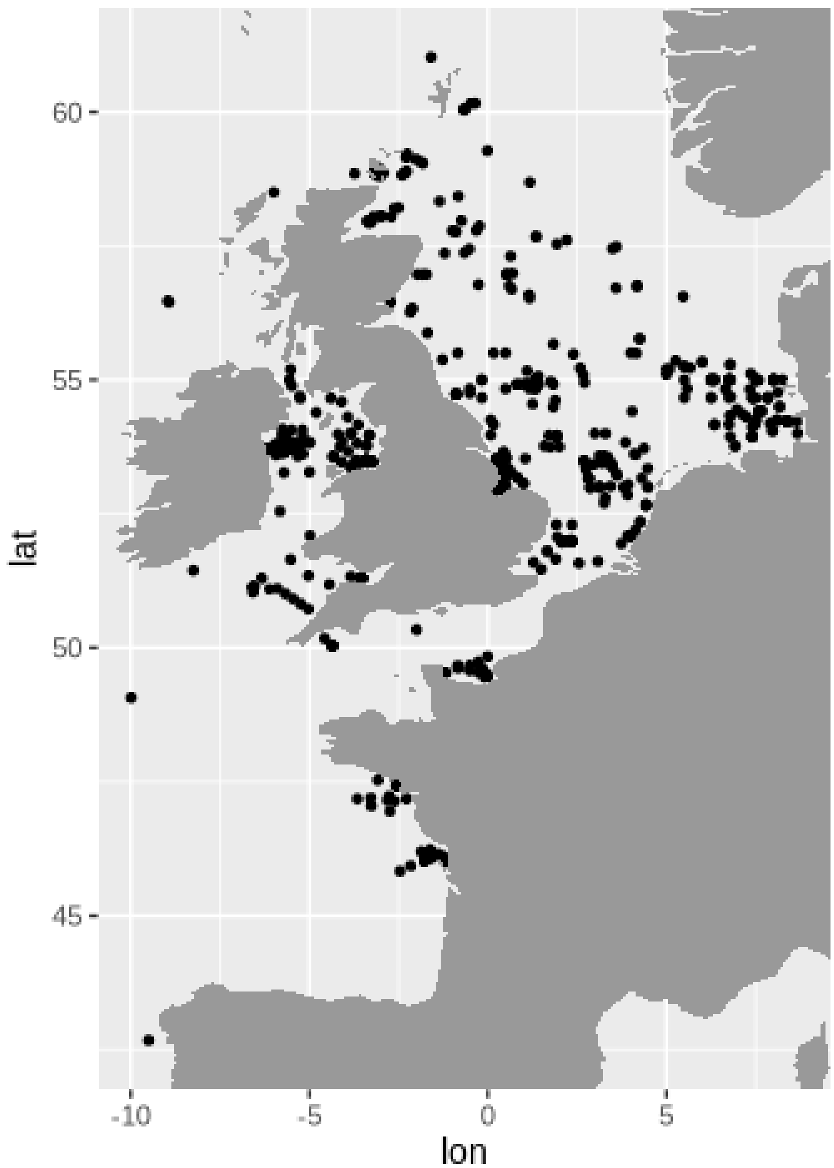


Figure 5.4: Locations of matchups between simulated and observed ammonium concentration near the bottom of the

water column.

The number of observations in each month ranged from 32 in December to 134 in February. Figure 5.5 below shows the distribution of observations in each month.

Figure 5.5 below shows the bias between the model and observational data for ammonium concentration. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

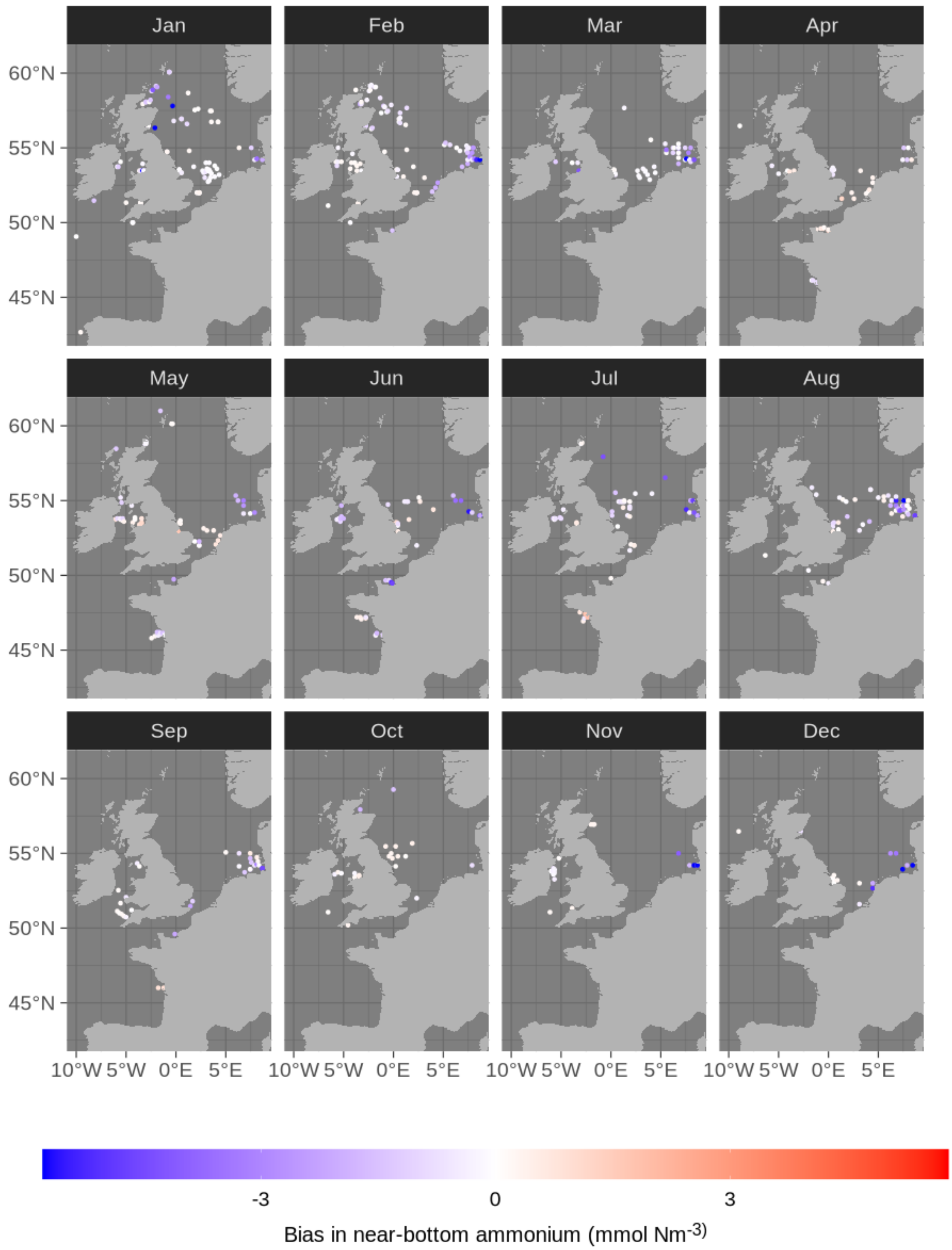


Figure 5.5: Bias in near-bottom ammonium concentration. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 5.6 and 5.7 show the distribution of near-bottom ammonium concentration observations in the model and observational datasets. This is shown for each month of the year (Figure 5.6) and for the entire year (Figure 5.7).

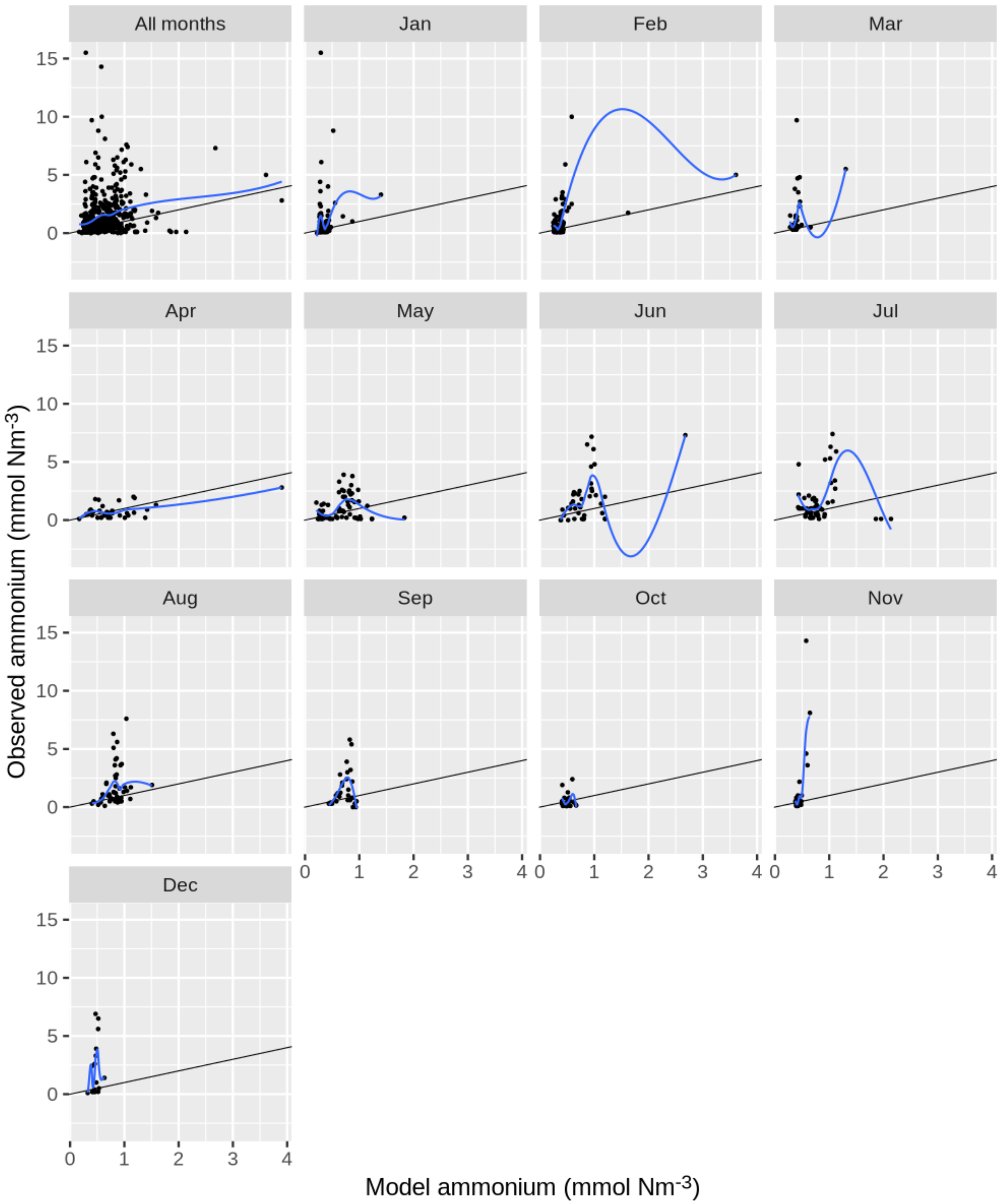


Figure 5.6: Simulated versus observed ammonium concentration near the bottom of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

5.4 Summary statistics for near-bottom ammonium

The overall ability of the model to predict the observed ammonium concentration was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	-0.73	1.94	0.22	919
Jan	-0.61	1.80	0.09	131
Feb	-0.49	1.27	0.38	134
Mar	-0.74	1.72	0.34	63
Apr	-0.29	0.91	0.22	61
May	-0.70	2.96	0.14	113
Jun	-1.24	2.38	0.40	69
Jul	-0.65	1.68	0.13	87
Aug	-0.91	1.85	0.19	77
Sep	-0.68	1.37	0.31	53
Oct	-0.18	0.66	-0.28	58
Nov	-1.61	2.95	0.67	41
Dec	-1.81	2.55	0.17	32

Table 5.3: Average bias (mmol Nm^{-3}) and root-mean square deviation (mmol Nm^{-3}) for the model's near-bottom ammonium concentration for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed ammonium concentration was performed. The modelled ammonium concentration was used as the independent variable and the observed ammonium concentration was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	1.16	0.63	0.05	0.00
Jan	1.34	0.58	0.01	0.37
Feb	1.51	0.47	0.15	0.00
Mar	4.04	-0.25	0.11	0.03
Apr	0.56	0.33	0.32	0.00
May	0.62	0.64	0.04	0.13
Jun	2.79	-0.38	0.28	0.00
Jul	0.41	1.25	0.01	0.56
Aug	2.29	-0.16	0.06	0.08
Sep	2.78	-0.54	0.09	0.10
Oct	-0.19	0.62	0.00	0.92
Nov	32.79	-13.12	0.54	0.00
Dec	9.46	-2.36	0.07	0.30

Table 5.4: Linear regression analysis of modelled and observed ammonium. The modelled ammonium concentration was used as the independent variable and the observed ammonium concentration was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

5.5 Depth-resolved comparisons of modelled and observed ammonium

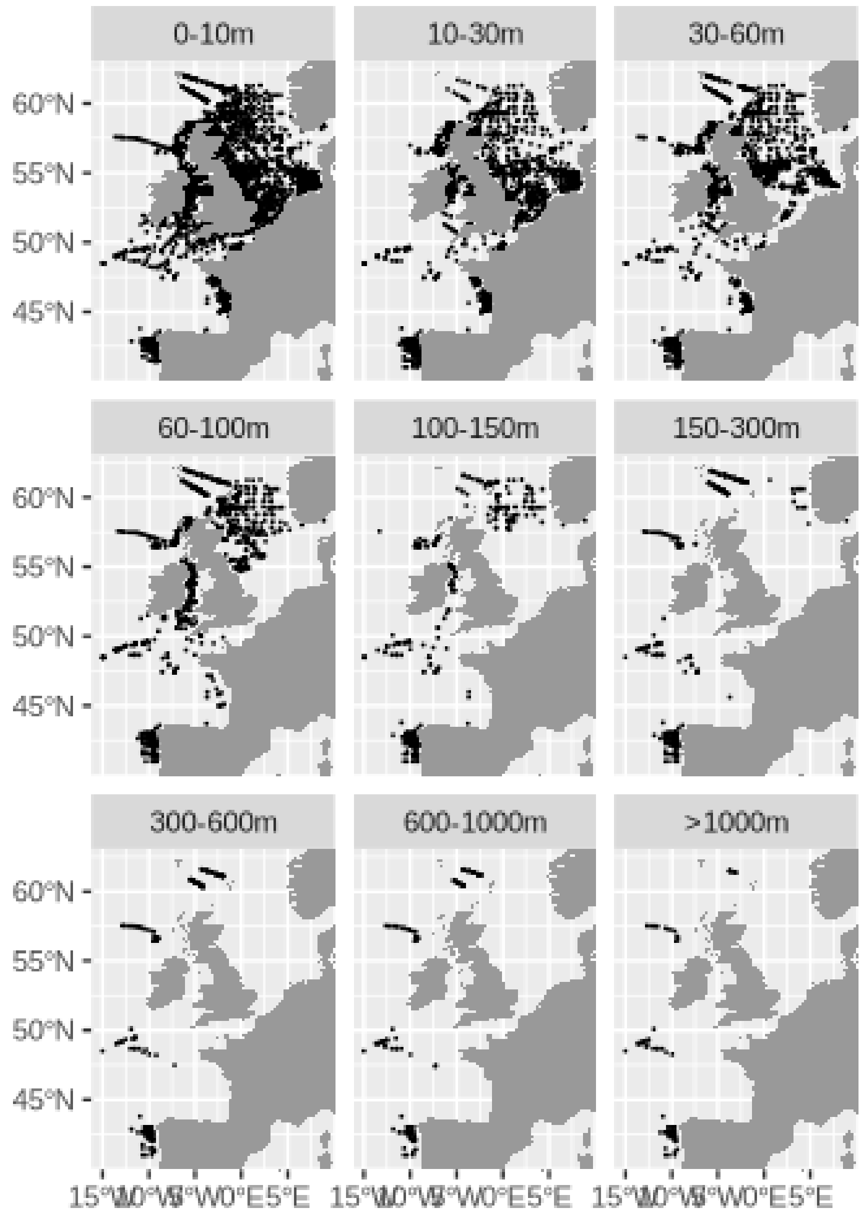


Figure 5.7: The geographic distribution of matchups between the model and observational ammonium concentration. The data has been binned into depth ranges. The depth ranges are 0-10m, 10-30m, 30-60m, 60-100m, 100-150m, 150-300m, 300-600m, 600-1000m, and >1000m. The number of observations in each depth range is shown in the tables below.

Depth	Bias	*r*	RMSD	Number of observations
0-150m	-0.47	0.25	1.28	13,745
0-10m	-0.47	0.21	1.42	6,151
10-30m	-0.41	0.27	1.17	3,313
30-60m	-0.56	0.39	1.22	2,391
60-100m	-0.50	0.21	1.15	1,449
100-150m	-0.19	0.33	0.64	441
150-300m	-0.43	-0.08	0.77	466
300-600m	-0.59	0.08	0.90	381
600-1000m	-0.61	0.22	0.92	312
>1000m	-0.36	0.29	0.69	219

Table 5.5: Average bias (mmol Nm^{-3}), root-mean square difference (RMSD) and correlation coefficient of modelled and observed ammonium concentration for different depth ranges. The bias is calculated as model-observation. The RMSD is the square root of the mean squared difference. The correlation coefficient is the Pearson correlation coefficient between the model and observed values.

5.6 Data Sources for validation of ammonium

ICES Data Portal, Dataset on Ocean HydroChemistry, Extracted March 3, 2023. ICES, Copenhagen

CHAPTER

SIX

SEA SURFACE AMMONIUM VALIDATION USING GRIDDED OBSERVATIONS FROM NSBC

We used version 1.1 of the **North Sea Biogeochemical Climatology** (NSBC) to validate **sea surface ammonium**. NSBC is a monthly climatology that covers the region 47°-65°N and 15°W-15°E. The data is made up of observations over the period 1960-2014. For validation purposes we only used the level 3 data, which a gridded monthly climatology at a spatial resolution of 1/4°. The data can be download from NSBC.

Matchup procedure: The model and observations were matched up as follows. First, the model dataset was cropped by a small amount to make sure cells close to the boundary were removed. The model was then regridded to the observational grid if the observational grid was coarser using nearest neighbour. Only grid cells with model and observational data were maintained. The following model output was used to compare with the observational values: **N4_n**.

6.1 Baseline climatologies of sea surface ammonium

Climatologies of model and observational sea surface ammonium concentration are shown in the figures below. The model climatology is calculated using the years **1986-2017**.

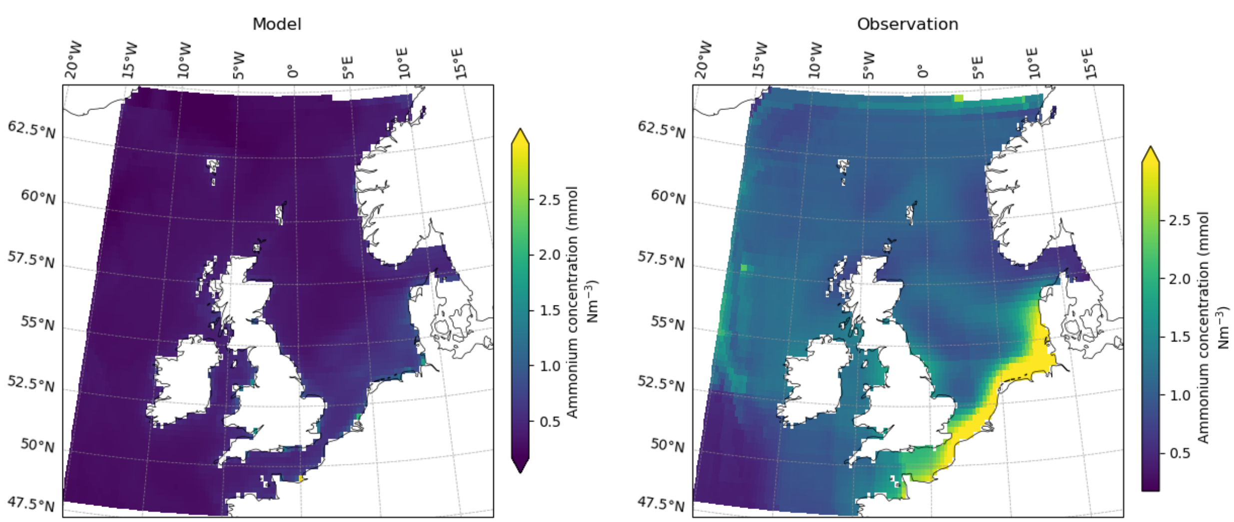


Figure 6.1: Annual average surface ammonium concentration from the model (1986-2017) and observations. Data is limited to the 2nd and 98th percentile of the combined model and observational data. Arrows indicate that values can exceed the colourbar limits.

6.2 Assessing model bias for surface ammonium concentration

Figure 6.2 shows the average bias of surface ammonium concentration simulated by the model. A positive bias indicates that the model overestimates the observation, while a negative bias indicates that the model overpredicts the observation.

The spatial average bias of surface ammonium concentration is -0.80 mmol Nm^{-3}. Overall, the model underestimates the observations in 99.8% of the model domain.

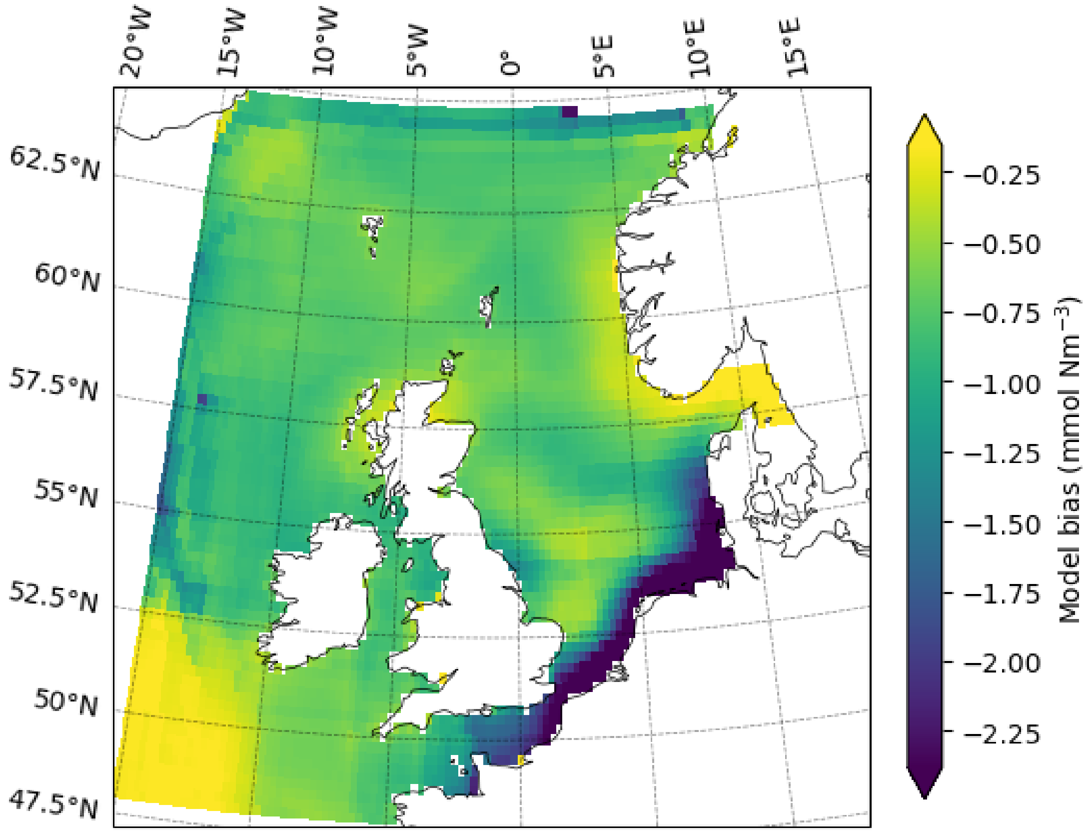


Figure 6.2: Bias of surface ammonium concentration from the model. A positive bias indicates that the model overestimates the observation. For clarity, the colourbar is limited to the 2nd and 98th percentile of the data.

6.3 Can the model reproduce seasonality of sea surface ammonium concentration?

The ability of the model to reproduce seasonality of sea surface ammonium concentration was assessed by comparing the modelled and observed seasonal cycle of ammonium concentration. First, we derive a monthly climatology for the model data. Then, we calculate the Pearson correlation coefficient between the modelled and observed ammonium concentration at each grid cell.

Note: we are only assessing the ability of the model to reproduce the ability of the model to reproduce seasonal changes, not long-term trends.

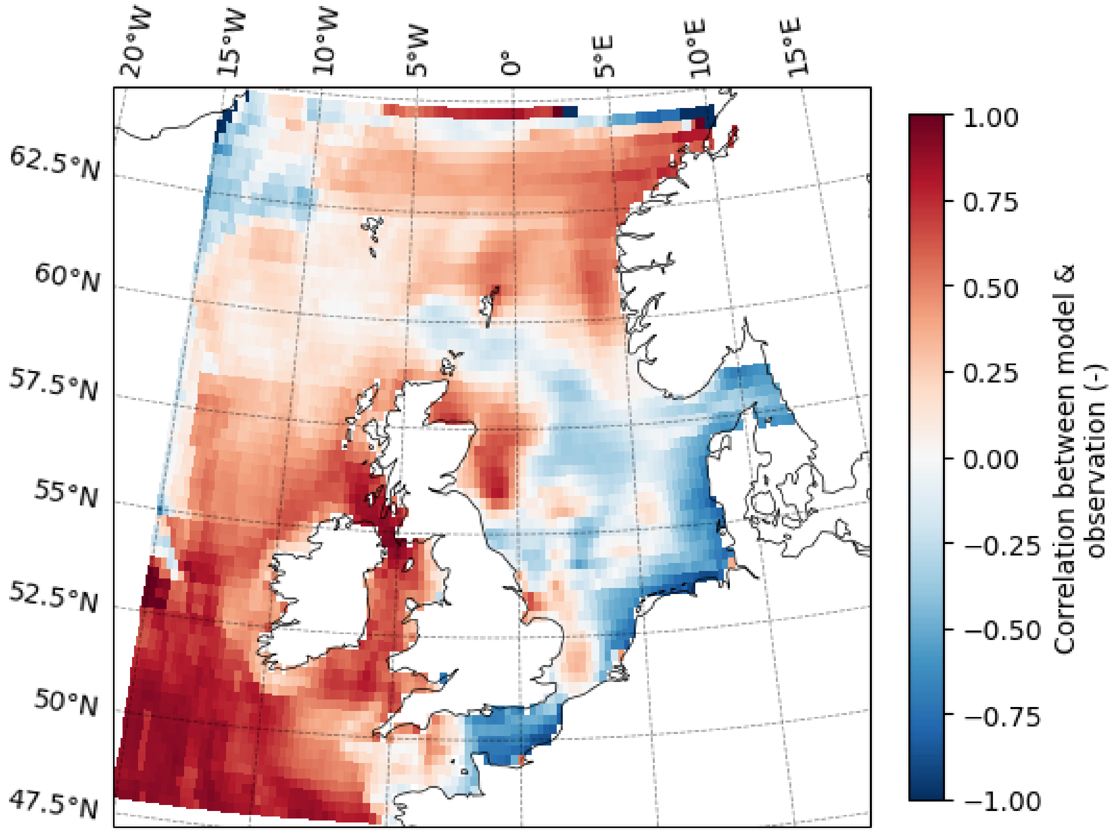
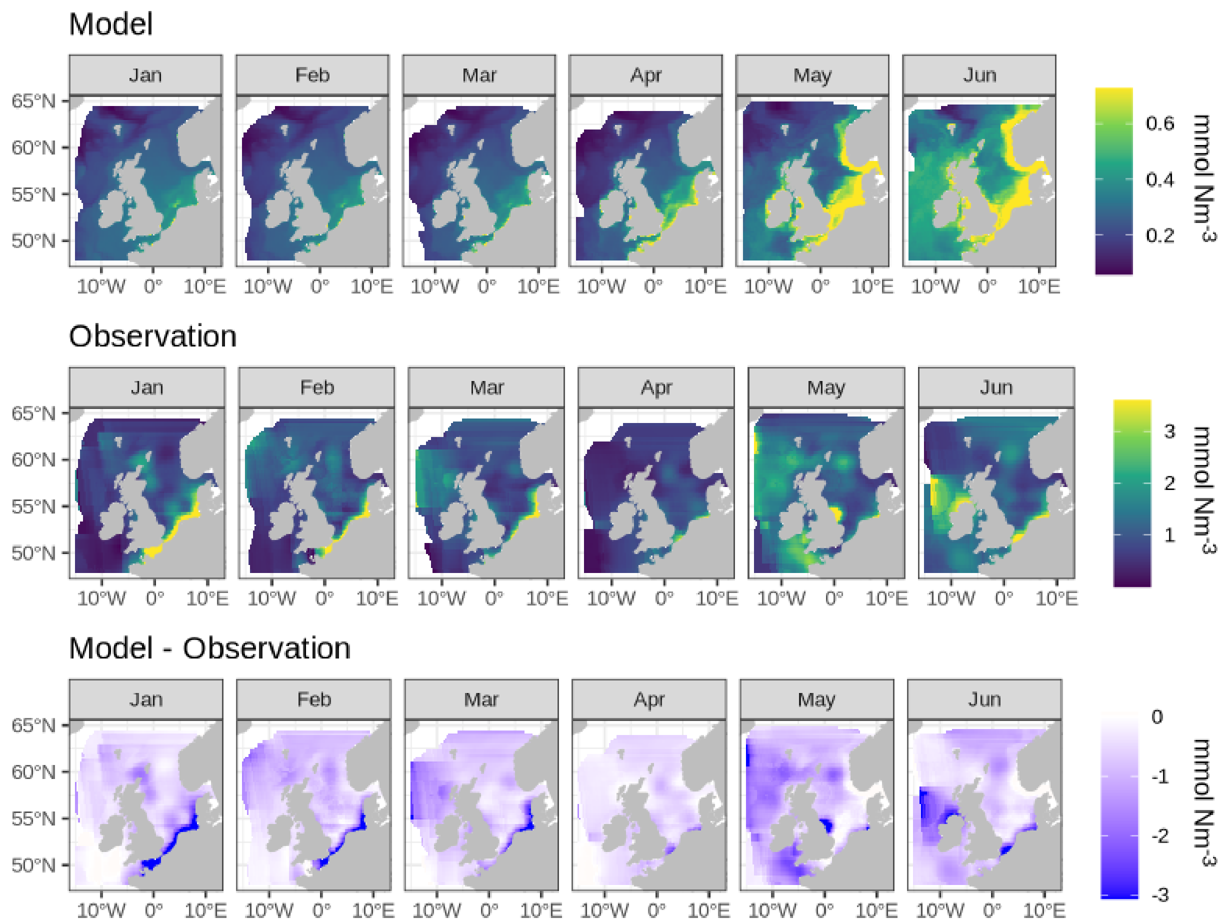


Figure 6.3: Seasonal temporal correlation between model and observations for surface ammonium concentration. This is the Pearson correlation coefficient between climatology monthly mean values in the model and observations.

The seasonal cycles of simulated and observed ammonium concentration are compared in Figure 6.4 below. This figure shows the model and observation average in each month of the year, and the differences between the two each month


Model
Jan
Feb
Mar
Apr
May
Jun
65°N
60°N
55°N
50°N
10°W 0° 10°E
0.6
0.4
0.2
mmol Nm-3
Observation
3
2
1
Model - Observation
0
-1
-2
-3

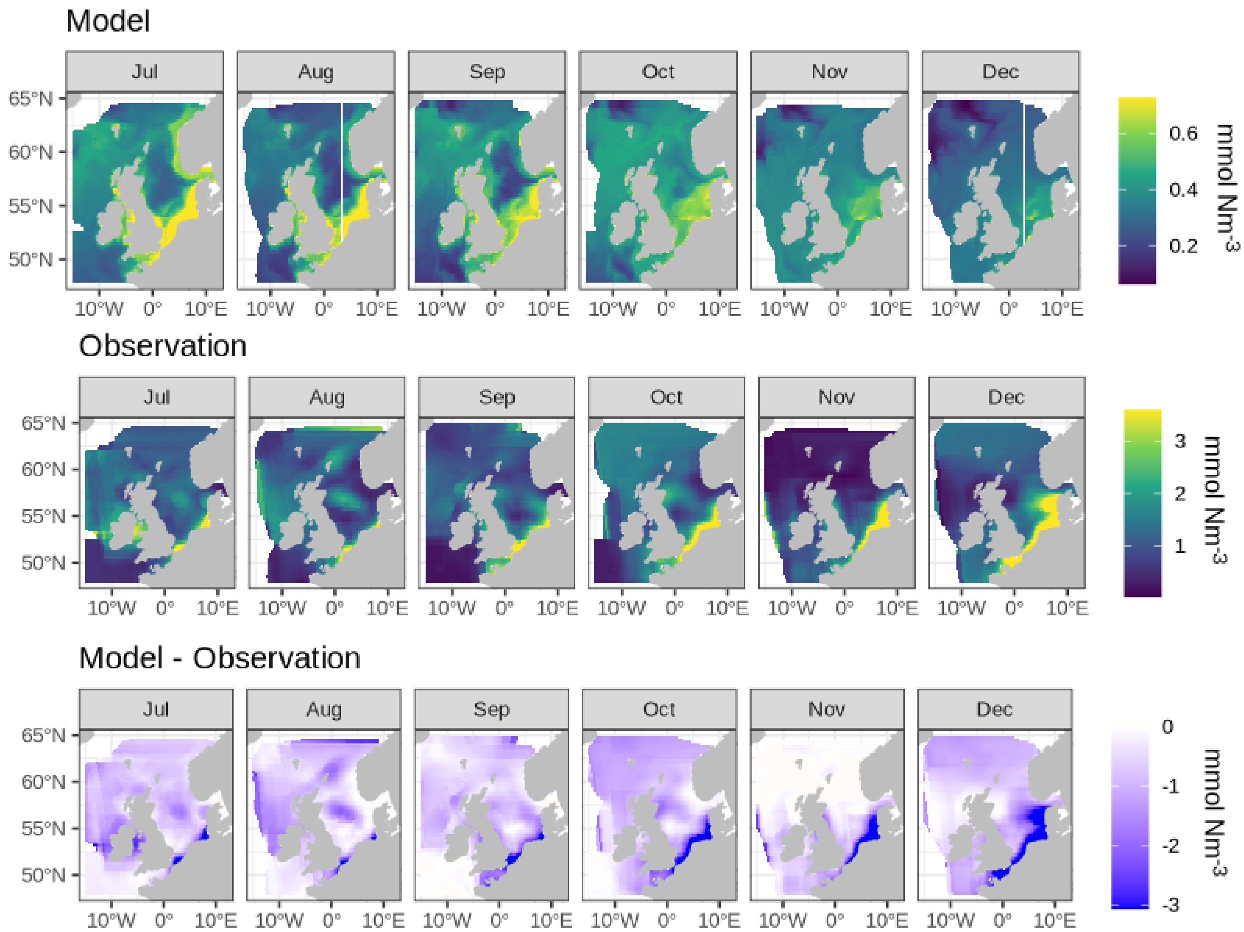


Figure 6.4: Monthly mean surface ammonium concentration for the model, observation and the difference between model and observations. For clarity, the maximum values are capped to the 98th percentiles.

6.4 Regional assessment of model performance for sea surface ammonium concentration

We assessed the regional performance of the model by comparing the model with observations in a number of regions. The regions considered are mapped below.

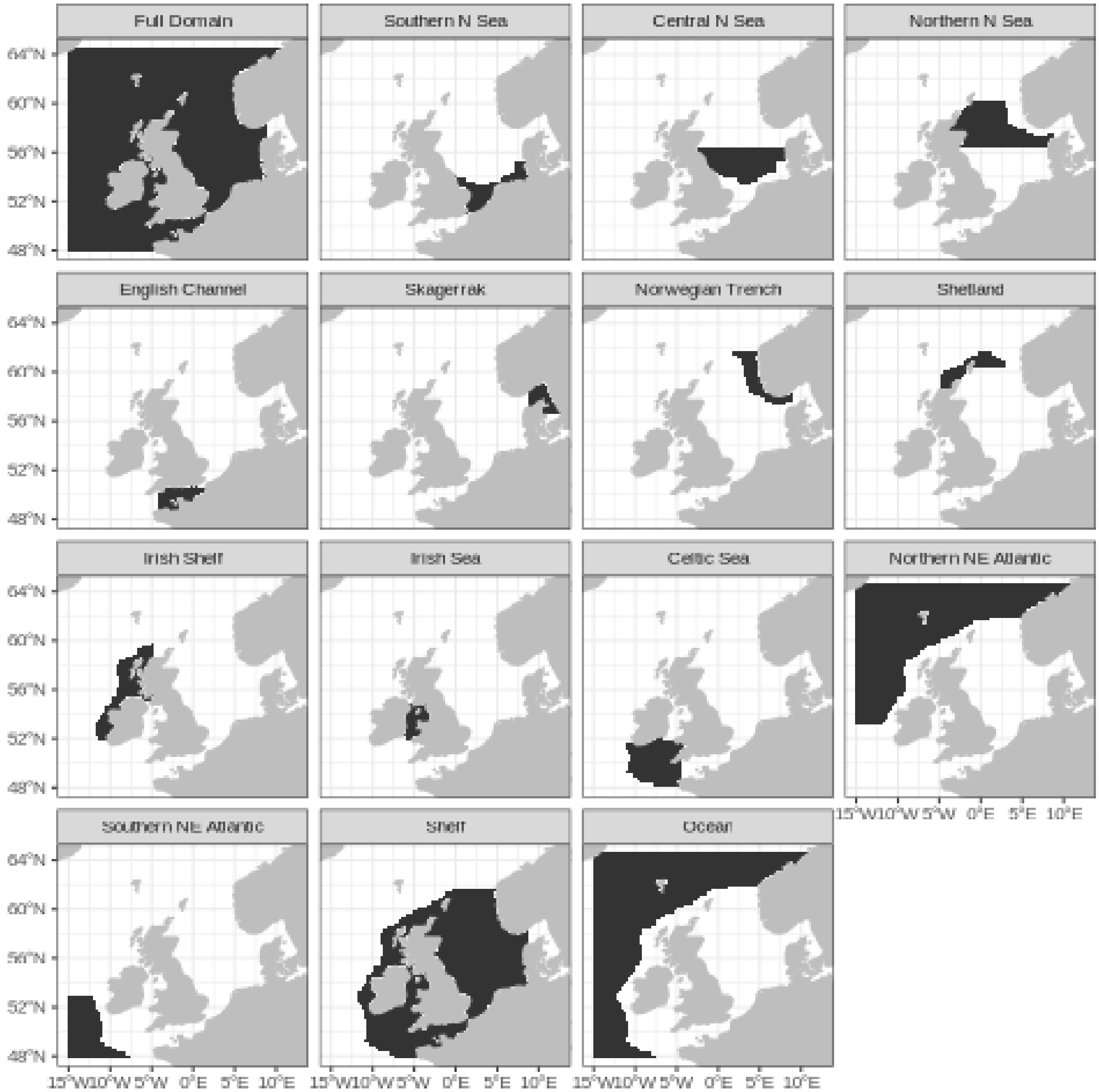


Figure 6.5: Regions used for validation of sea surface ammonium concentration.

Time series were constructed comparing the monthly mean of the spatial average sea surface ammonium concentration in each region. The spatial average was calculated using the mean of all grid cells within each region, accounting for grid cell area.

Figure 6.6: Seasonal cycle of sea surface ammonium concentration for model and observations for each region. The spatial average is taken over the region.

The table below shows the average bias of sea surface ammonium concentration in each region. The bias is calculated as the modelled value minus the observed value. A positive bias indicates that the model overestimates the observed value,

while a negative bias indicates that the model underestimates the observed value.

Region	Temporal correlation	Bias	RMSD
Full Domain	0.24	-0.81	1.13
Celtic Sea	0.52	-0.63	0.83
Central North Sea	-0.10	-0.83	1.14
Channel	-0.26	-1.44	1.84
Irish Sea	0.57	-0.82	0.97
Irish Shelf	0.46	-0.74	0.93
Northern North East Atlantic	0.21	-0.80	0.96
Northern North Sea	0.00	-0.73	0.91
Norwegian Trench	0.18	-0.40	0.55
Ocean	0.32	-0.71	0.91
Shelf	0.15	-0.90	1.30
Shetland	0.20	-0.74	0.86
Skagerrak	-0.49	-0.13	0.54
Southern North East Atlantic	0.82	-0.33	0.46
Southern North Sea	-0.27	-2.28	3.03

Table 6.1: Summary of performance of the model sea surface ammonium concentration in each region. The bias (mmol Nm-3) column represents the spatial average of the annual mean modelled value minus the observed value. The temporal correlation column represents the spatial mean of the temporal correlation between the model and observations per grid cell.

6.5 Can the model reproduce spatial patterns of sea surface ammonium concentration?

The ability of the model to reproduce spatial patterns of sea surface ammonium concentration was assessed by comparing the modelled and observed ammonium concentration at each grid cell. We calculated the Pearson correlation coefficient between the modelled and observed ammonium concentration at each grid cell.

This was carried out monthly and using the annual mean in each grid cell

Time period	*r*
Annual mean	0.61
Jan	0.49
Feb	0.37
Mar	0.34
Apr	0.66
May	-0.06
Jun	0.27
Jul	0.47
Aug	0.21
Sep	0.60
Oct	0.48
Nov	0.53
Dec	0.45

Table 6.2: Pearson correlation coefficient between modelled and observed sea surface ammonium concentration at each grid cell. The correlation was calculated monthly and using the annual mean in each grid cell.

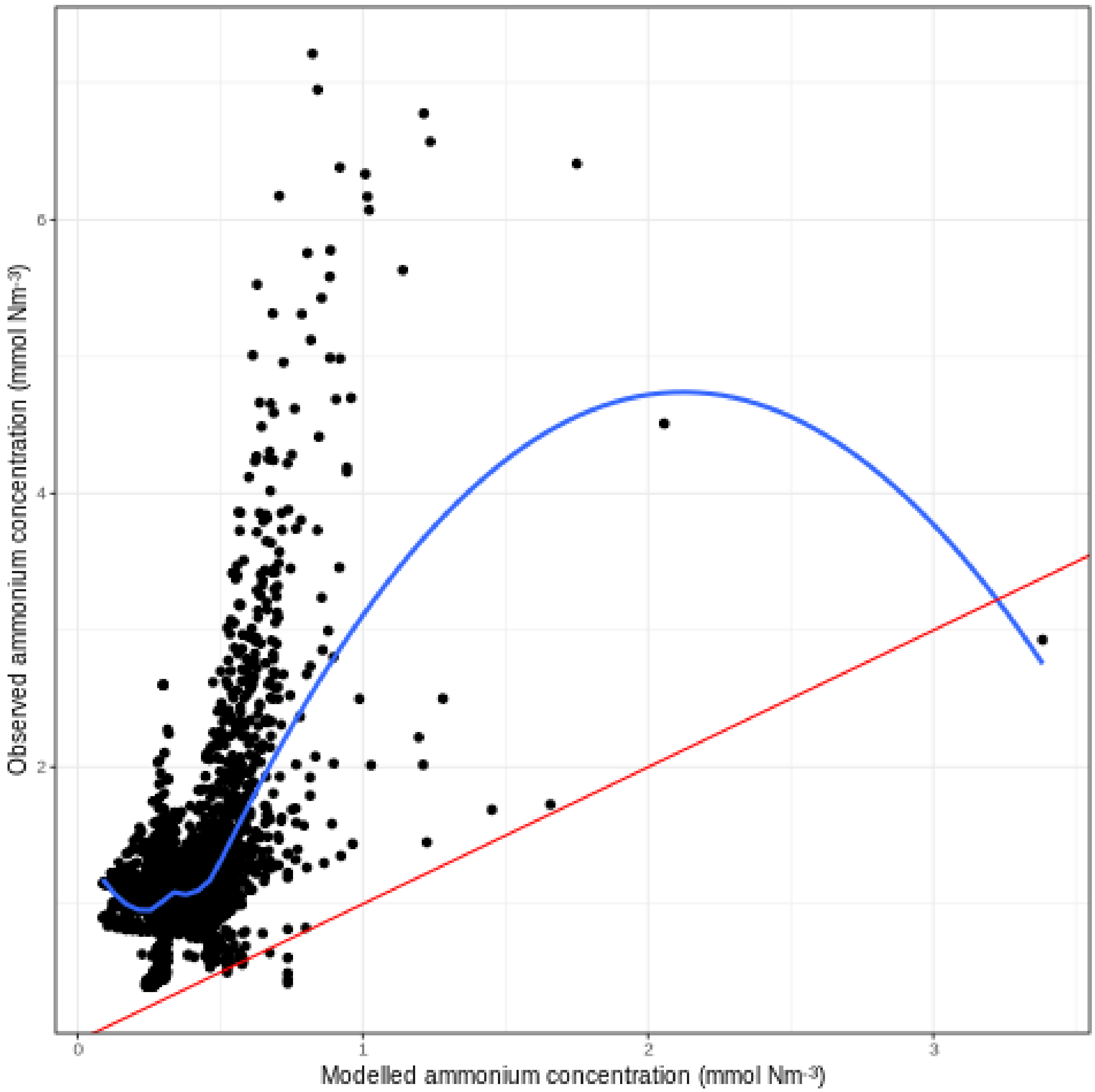


Figure 6.7: Scatter plot of modelled and observed annual average sea surface ammonium concentration in grid cells. The red line is the 1:1 line. The blue line is the linear regression of the modelled and observed values. The slope of the blue line is the slope of the regression line.

6.6 Data Sources for validation of ammonium concentration

Hinrichs, Iris; Gouretski, Viktor; Paetsch, Johannes; Emeis, Kay; Stammer, Detlef (2017). North Sea Biogeochemical Climatology (Version 1.1).

URL: https://www.cen.uni-hamburg.de/en/icdc/data/ocean/nsbc.html

CHAPTER

SEVEN

VALIDATION OF BENTHIC MACROBENTHOS BIOMASS USING POINT OBSERVATIONS

7.1 Performance of model biomass of macrobenthos in the sediment layer

Biomass data for macrobenthos was downloaded from the North Sea Benthos Survey 1986. The model output was matched up with the observational data with model output from the years **1986 to 2017**.

In total there were 171 values extracted from the observational database. The map below shows the locations of the matched up data for biomass of macrobenthos.

The following model output was used to compare with observational values: **Y2_c+Y3_c**.

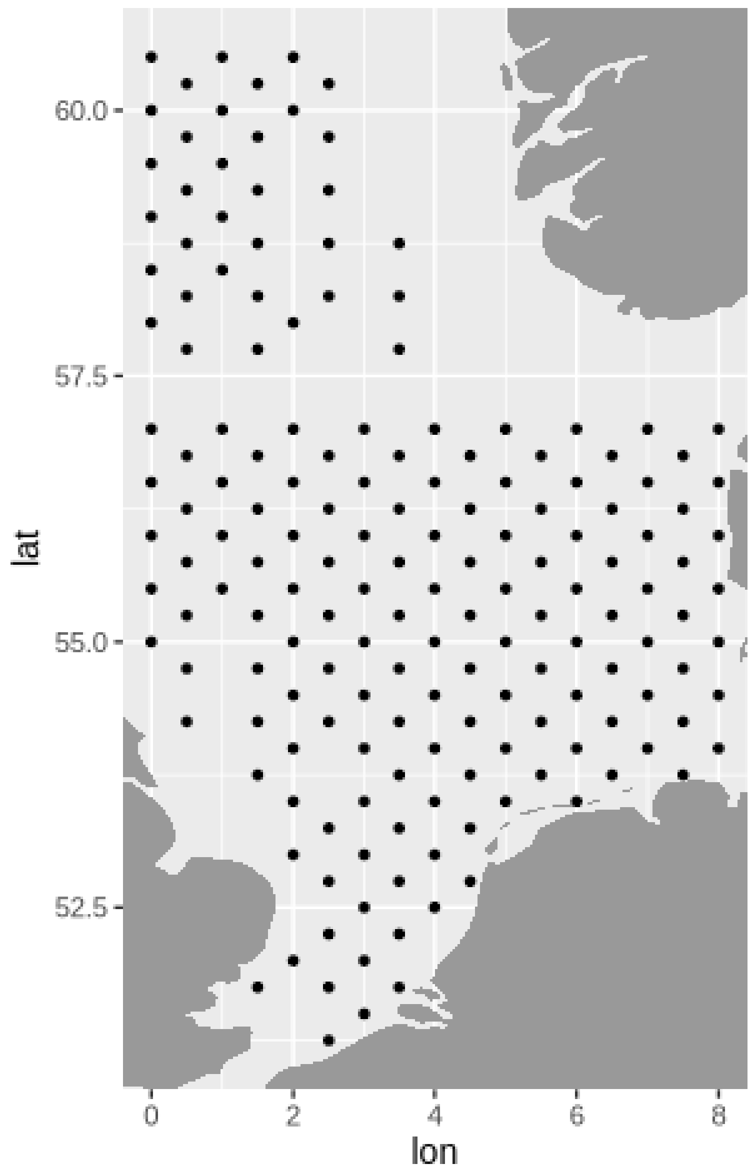


Figure 7.1: Locations of matchups between simulated and observed biomass of macrobenthos on the seafloor. The observational data is from the North Sea Benthos Survey (1986).

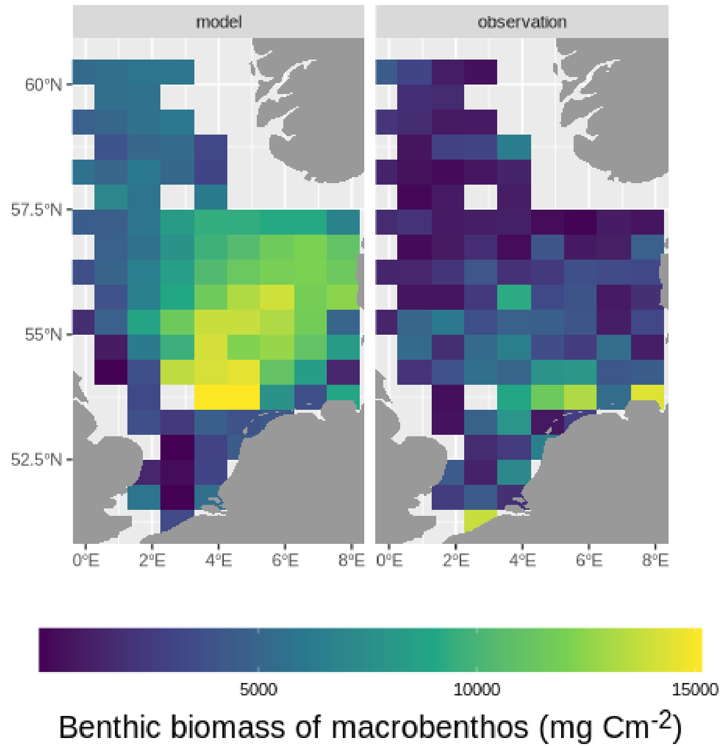


Figure 7.2: Map of average benthic biomass of macrobenthos in the model and observational datasets.

Figure 7.3 below shows the bias between the model and observational data for biomass of macrobenthos. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

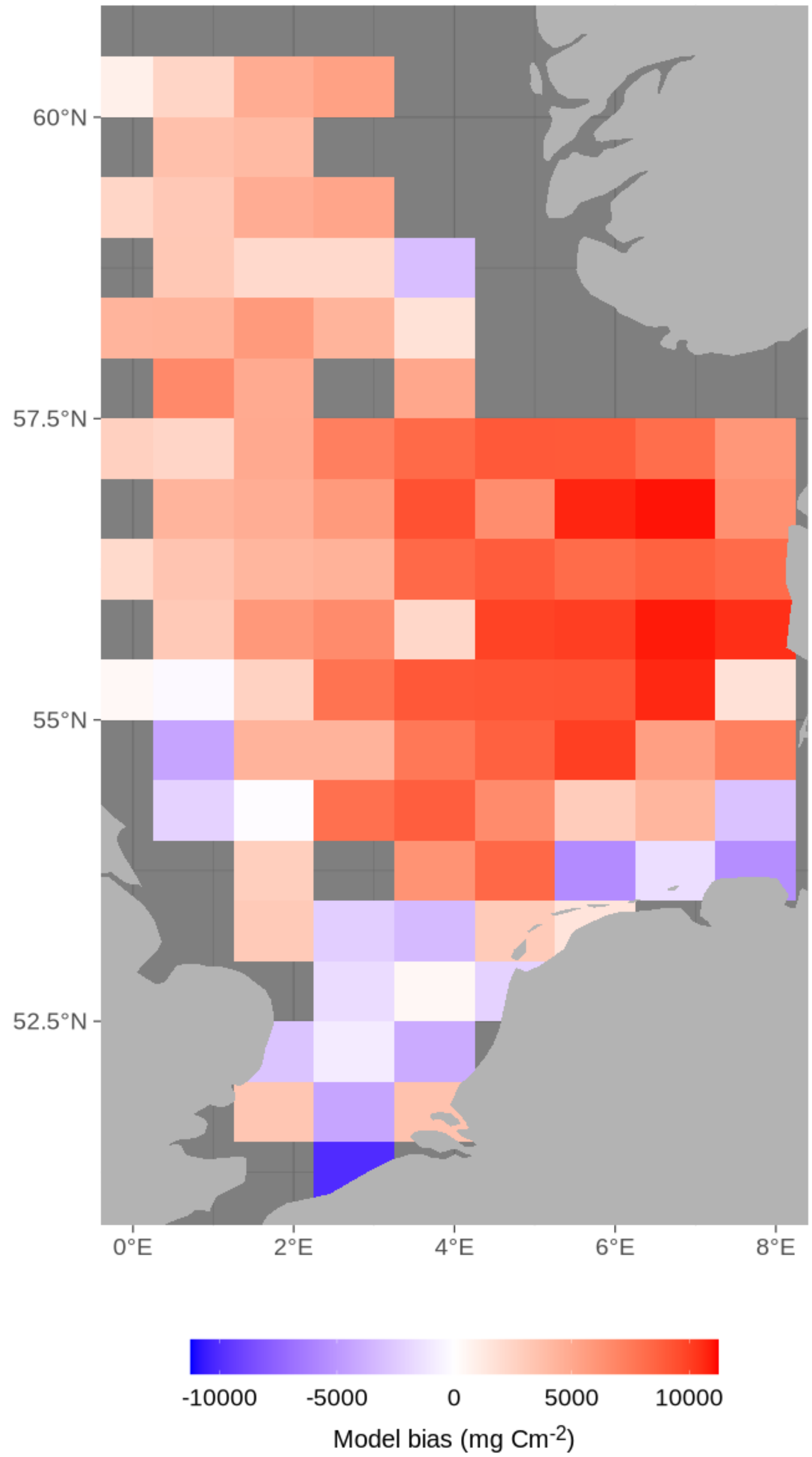


Figure 7.3: Bias in benthic biomass of macrobenthos. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale.

Figures 7.4 and 7.5 show the distribution of benthic biomass of macrobenthos observations in the model and observational datasets. This is shown for each month of the year (Figure 7.4) and for the entire year (Figure 7.5).

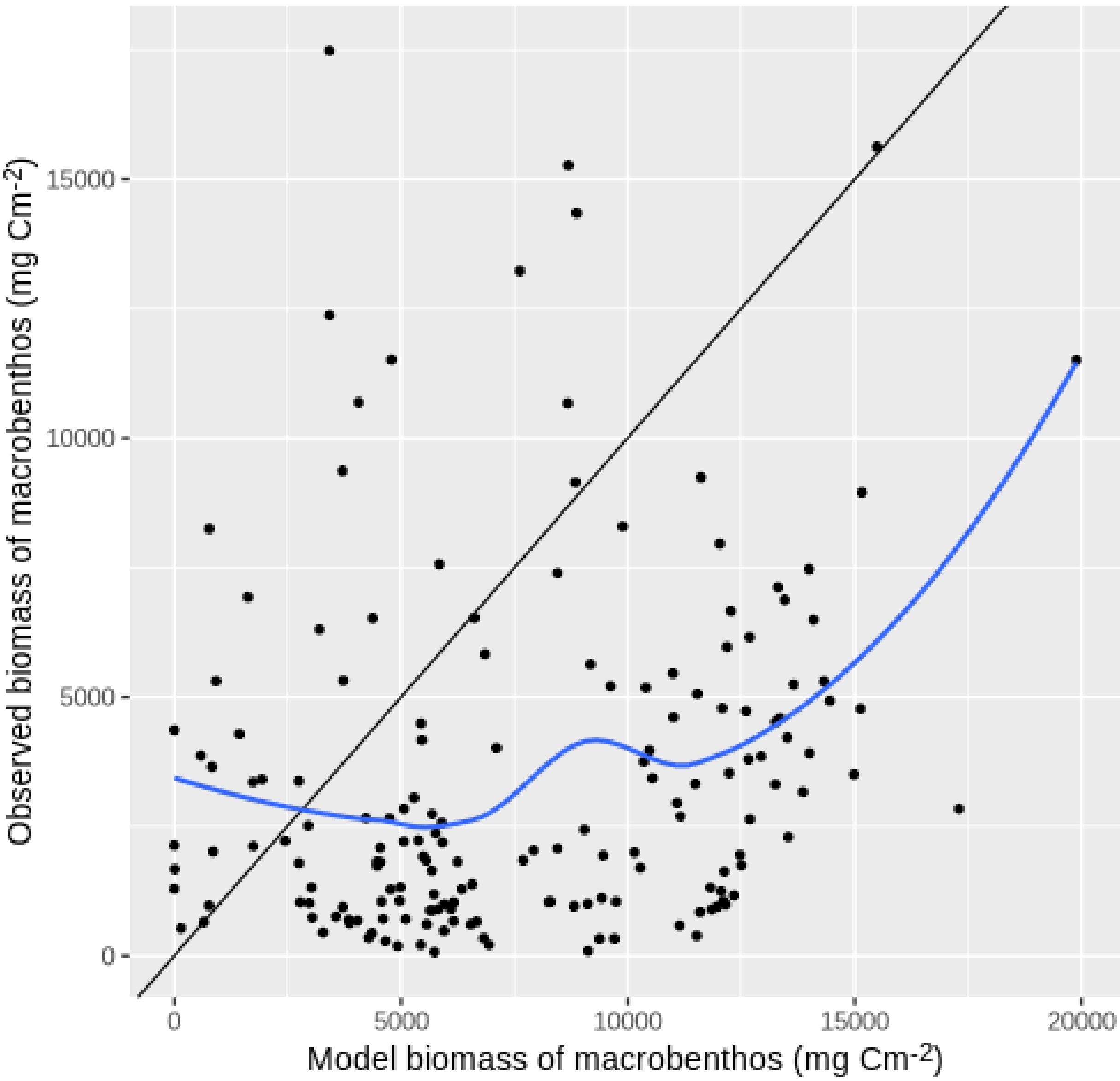


Figure 7.4: Model vs observed biomass of macrobenthos for benthic values. The line is a generalized additive model (GAM) fit to the data. The shaded area is the 95% confidence interval of the GAM fit.

7.2 Summary statistics for biomass of macrobenthos in the sediment layer

The overall ability of the model to predict the observed biomass of macrobenthos was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	7592.25	8704.36	0.22	171

Table 7.1: Average bias (mg Cm^{-2}) and root-mean square deviation (mg Cm^{-2}) for the model's benthic biomass of macrobenthos for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed biomass of macrobenthos was performed. The modelled biomass of macrobenthos was used as the independent variable and the observed biomass of macrobenthos was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.17	2169.78	0.05	0.00

Table 7.2: Linear regression analysis of modelled and observed biomass of macrobenthos. The modelled biomass of macrobenthos was used as the independent variable and the observed biomass of macrobenthos was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

7.3 Data Sources for validation of biomass of macrobenthos

URL: https://www.vliz.be/vmdcdata/nsbs/about.php

CHAPTER

EIGHT

VALIDATION OF BENTHIC CARBON USING POINT OBSERVATIONS

8.1 Performance of model carbon in the sediment layer

Carbon data was compiled from multiple sources The model output was matched up with the observational data with model output from the years **1986 to 2017**.

In total there were 289 values extracted from the observational database. The map below shows the locations of the matched up data for carbon.

The following model output was used to compare with observational values: **1e-6*Q6_c * (1 - exp(-0.1 / Q6_pen_depth_c)) + Q7_c * (1 - exp(-0.1 / Q7_pen_depth_c))/0.1**.

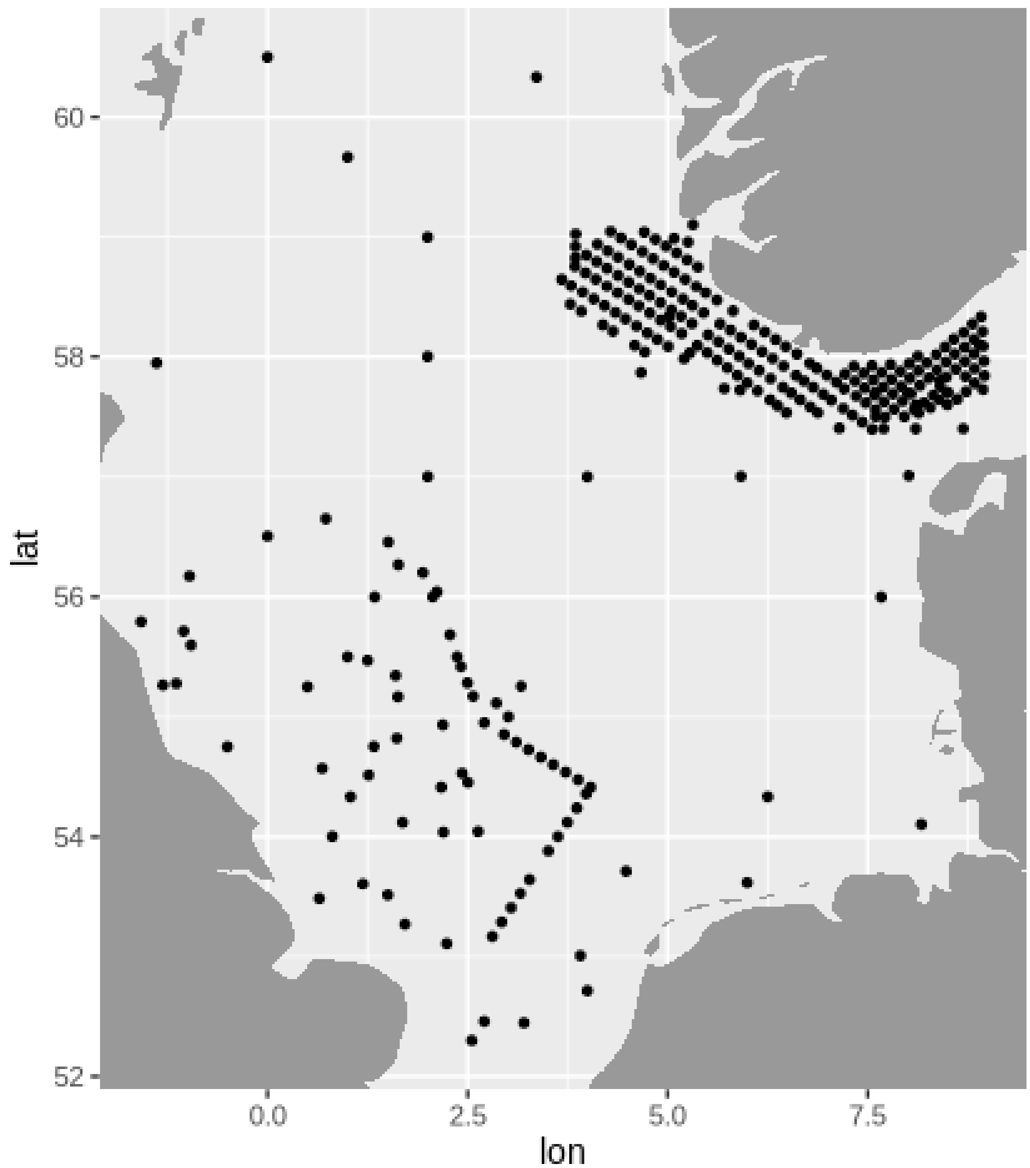


Figure 8.1: Locations of matchups between simulated and observed carbon on the seafloor. The observational data is from Diesing et al. (2021).

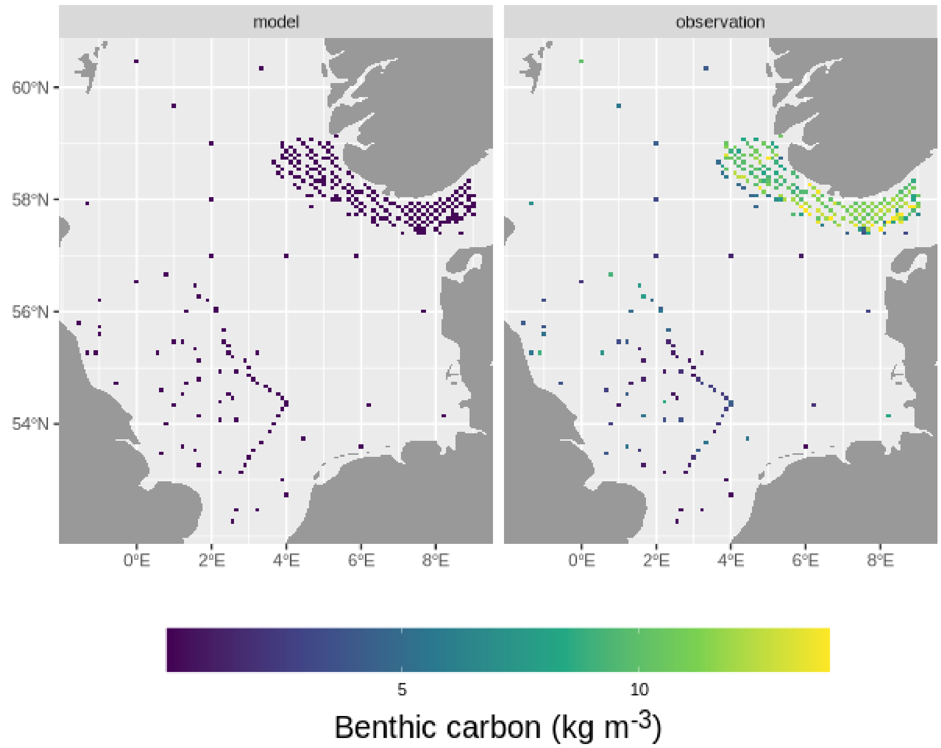


Figure 8.2: Map of average benthic carbon in the model and observational datasets.

Figure 8.3 below shows the bias between the model and observational data for carbon. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

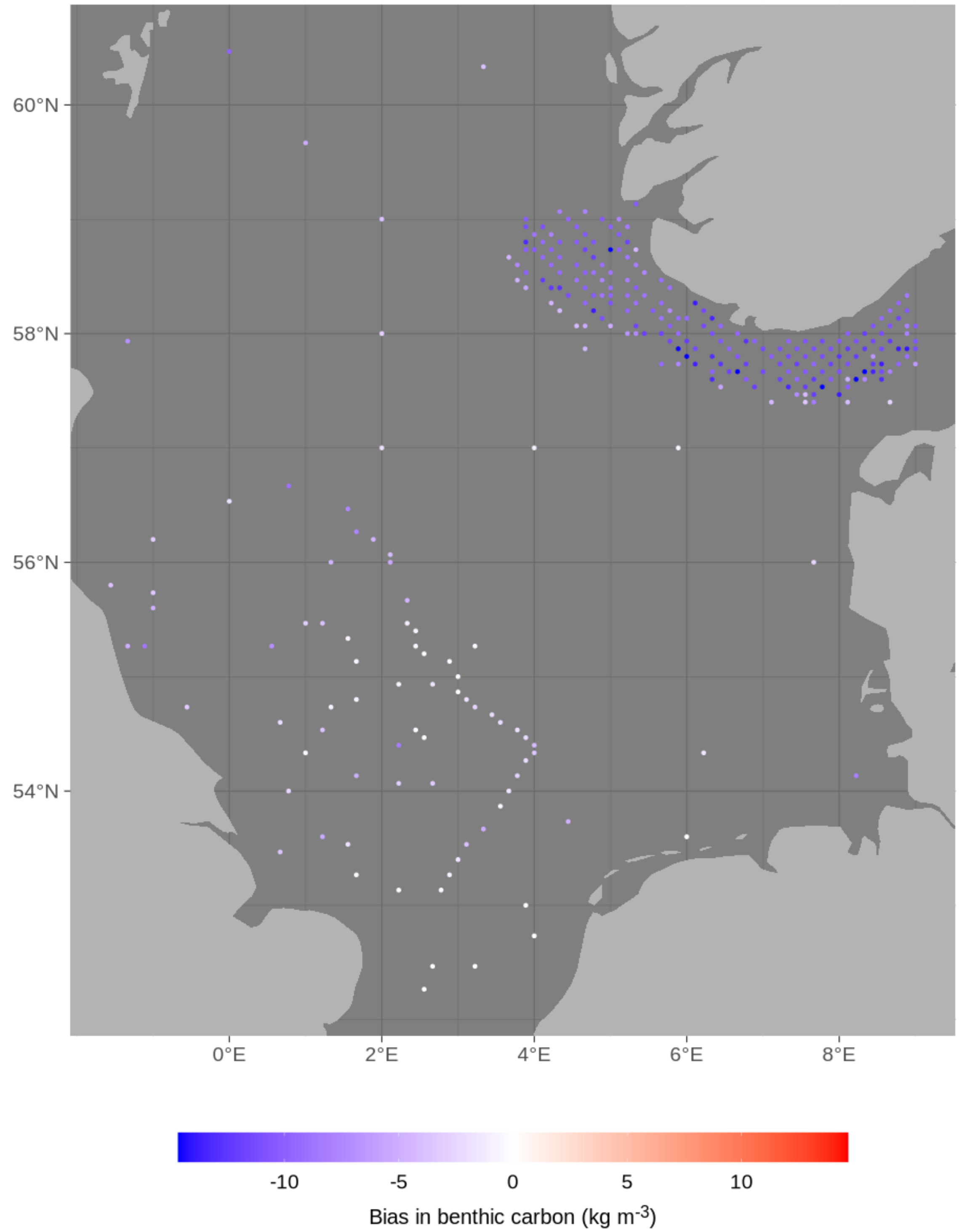


Figure 8.3: Bias in benthic carbon. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale.

Figures 8.4 and 8.5 show the distribution of benthic carbon observations in the model and observational datasets. This is shown for each month of the year (Figure 8.4) and for the entire year (Figure 8.5).

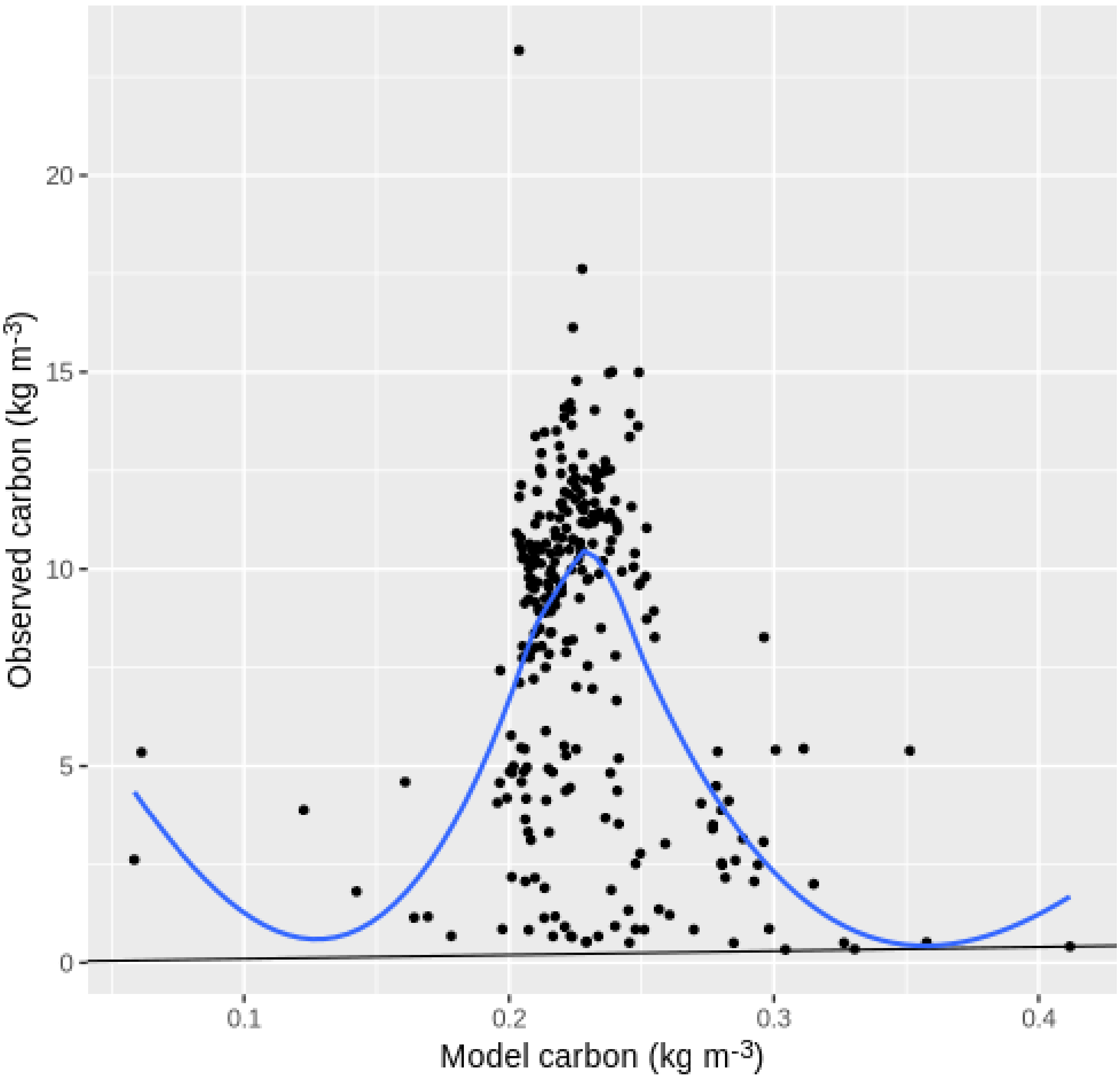


Figure 8.4: Model vs observed carbon for benthic values. The line is a generalized additive model (GAM) fit to the data. The shaded area is the 95% confidence interval of the GAM fit.

8.2 Summary statistics for carbon in the sediment layer

The overall ability of the model to predict the observed carbon was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	-8.00	9.03	-0.20	289

Table 8.1: Average bias (kg m^{-3}) and root-mean square deviation (kg m^{-3}) for the model's benthic carbon for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed carbon was performed. The modelled carbon was used as the independent variable and the observed carbon was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	-24.19	13.73	0.04	0.00

Table 8.2: Linear regression analysis of modelled and observed carbon. The modelled carbon was used as the independent variable and the observed carbon was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

8.3 Data Sources for validation of carbon

Diesing, Markus, Terje Thorsnes, and Lilja Rún Bjarnadóttir. "Organic carbon densities and accumulation rates in surface sediments of the North Sea and Skagerrak." Biogeosciences 18.6 (2021): 2139-2160.

CHAPTER

NINE

VALIDATION OF CHLOROPHYLL USING POINT OBSERVATIONS

9.1 Performance of model sea surface chlorophyll

Values from the **top 5m** of the water column were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 12,814 values extracted from the observational database. The map below shows the locations of the matched up data for chlorophyll concentration.

The following model output was used to compare with observational values: **P1_Chl+P2_Chl+P3_Chl+P4_Chl**.

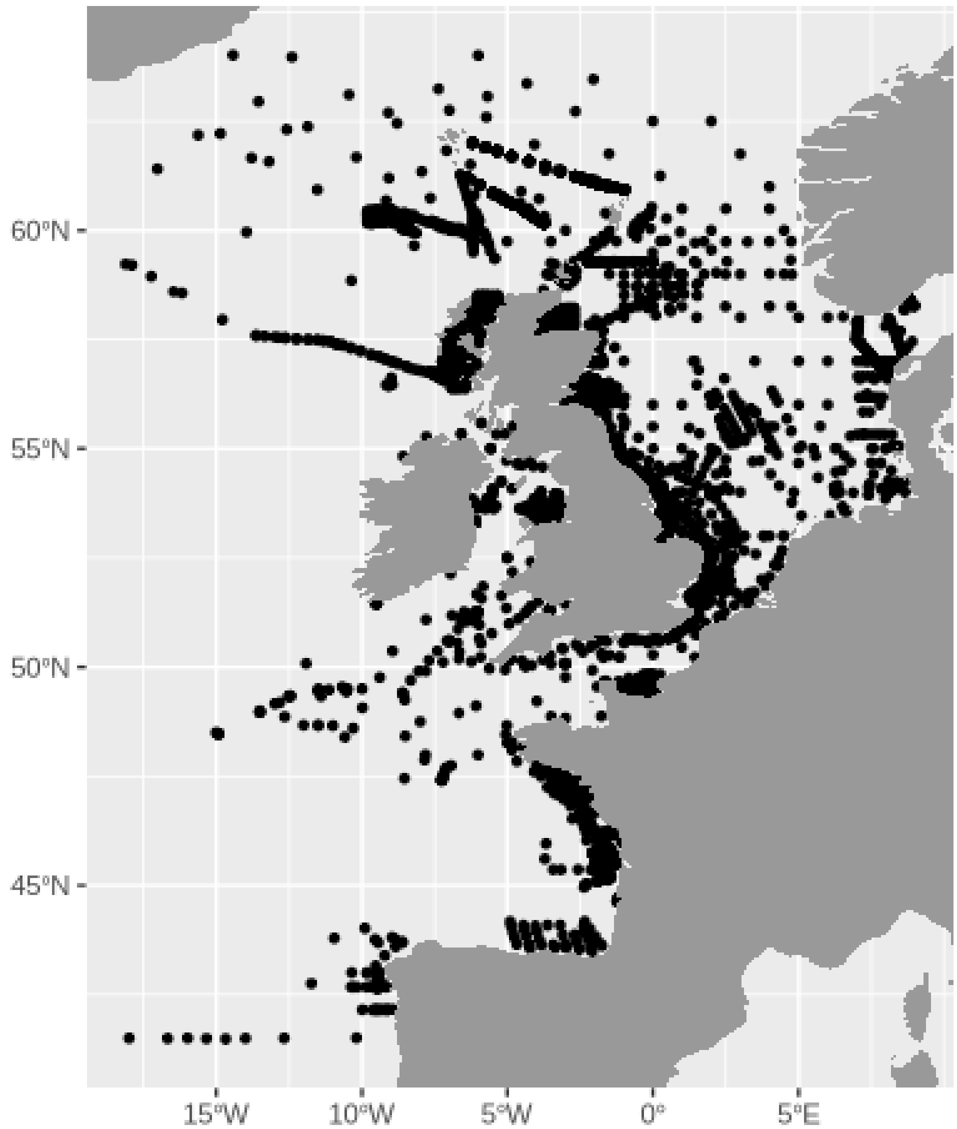


Figure 9.1: Locations of matchups between simulated and observed chlorophyll concentration in the top 5m of the water column.

The number of observations in each month ranged from 550 in November to 1,852 in May. Figure 9.2 below shows the distribution of observations in each month.

Figure 9.2 below shows the bias between the model and observational data for chlorophyll concentration. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

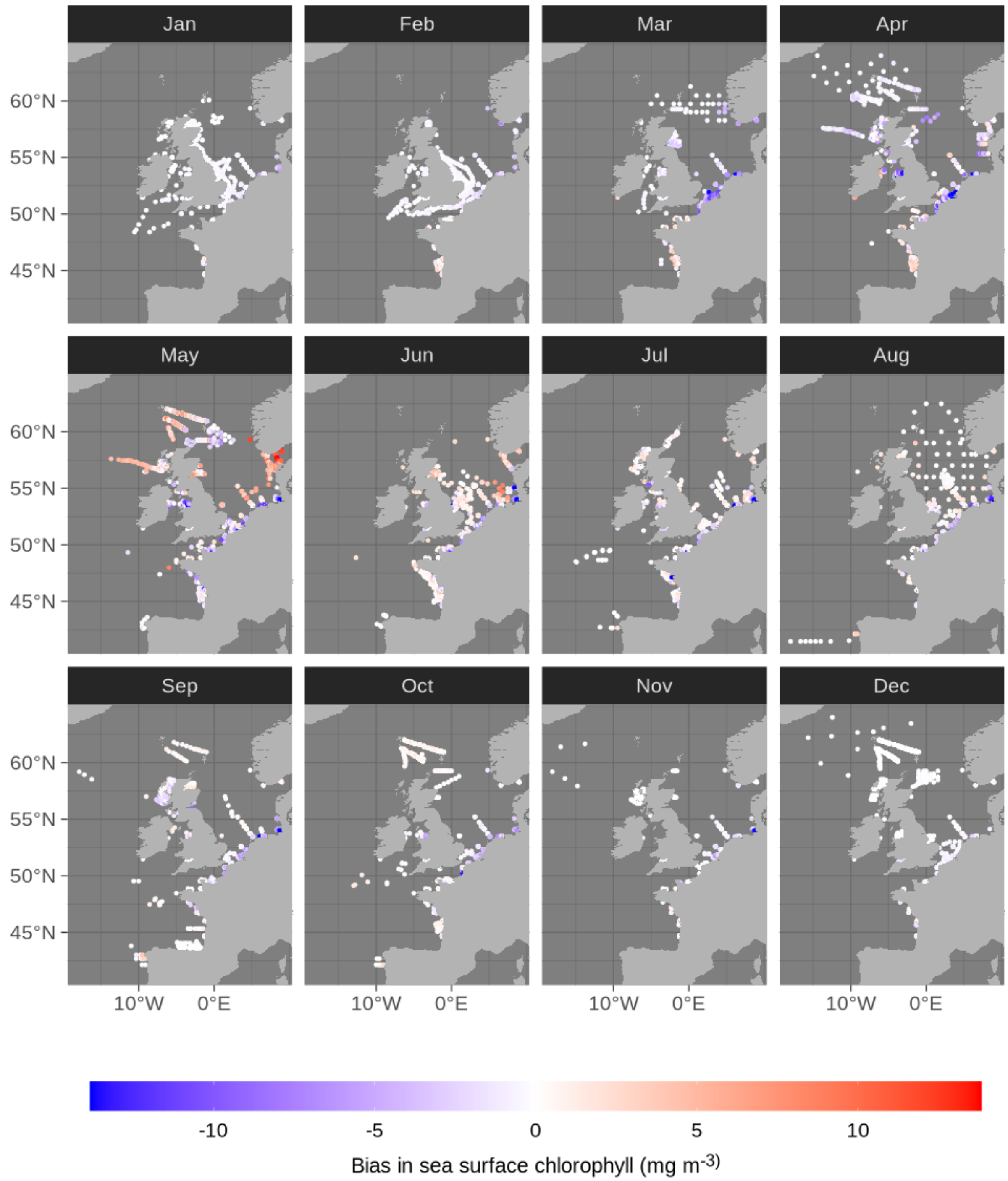


Figure 9.2: Bias in sea surface chlorophyll concentration. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 9.3 and 9.4 show the distribution of sea surface chlorophyll concentration observations in the model and observational datasets. This is shown for each month of the year (Figure 9.3) and for the entire year (Figure 9.4).

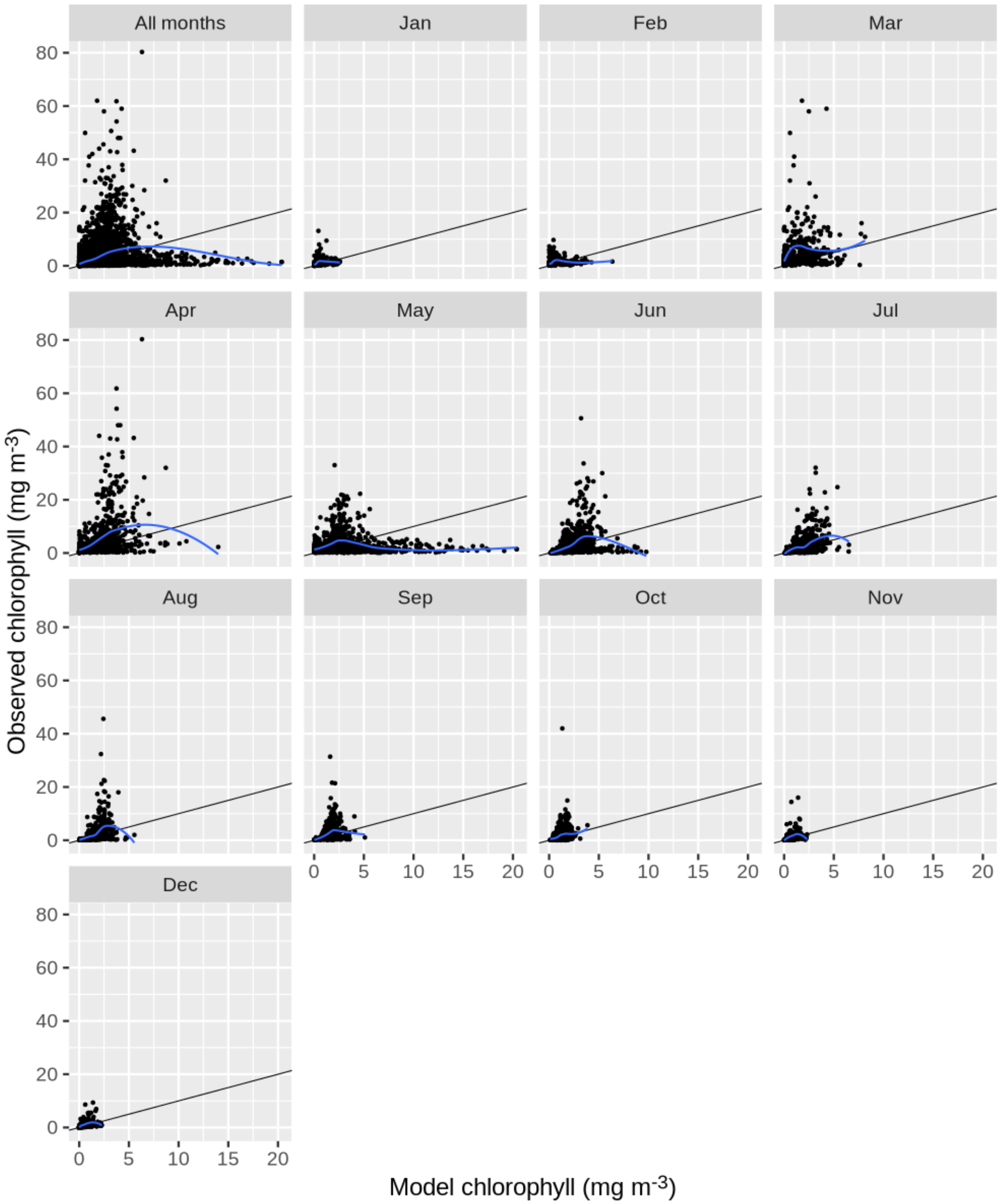


Figure 9.3: Simulated versus observed chlorophyll concentration in the top 5m of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

9.2 Summary statistics for sea surface chlorophyll

The overall ability of the model to predict the observed chlorophyll concentration was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	-0.80	4.45	0.21	12,814
Jan	-0.29	1.11	0.30	880
Feb	-0.48	1.58	0.17	825
Mar	-2.50	6.69	0.12	772
Apr	-2.39	7.53	0.32	1,509
May	-0.11	5.25	-0.06	1,852
Jun	-0.57	5.45	0.10	1,354
Jul	-0.75	3.58	0.32	1,126
Aug	-0.82	3.51	0.34	1,000
Sep	-0.71	2.42	0.37	1,091
Oct	-0.43	2.22	0.17	920
Nov	-0.23	1.42	0.18	550
Dec	-0.16	0.89	0.47	935

Table 9.1: Average bias (mg m^{-3}) and root-mean square deviation (mg m^{-3}) for the model's sea surface chlorophyll concentration for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed chlorophyll concentration was performed. The modelled chlorophyll concentration was used as the independent variable and the observed chlorophyll concentration was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.62	1.56	0.06	0.00
Jan	0.61	0.64	0.09	0.00
Feb	0.29	0.96	0.04	0.00
Mar	0.81	3.52	0.03	0.00
Apr	1.55	1.77	0.13	0.00
May	-0.05	3.12	0.00	0.14
Jun	0.79	1.42	0.05	0.00
Jul	1.45	-0.01	0.16	0.00
Aug	1.66	-0.00	0.13	0.00
Sep	1.16	0.73	0.08	0.00
Oct	1.07	0.43	0.05	0.00
Nov	0.62	0.87	0.05	0.00
Dec	1.08	0.24	0.31	0.00

Table 9.2: Linear regression analysis of modelled and observed chlorophyll. The modelled chlorophyll concentration was used as the independent variable and the observed chlorophyll concentration was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

9.3 Performance of model near-bottom chlorophyll

Near-bottom values of chlorophyll were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The near-bottom was defined as observations **within 2m of the seabed**. This was interpolated to the observational grid using the $GEBCO_2$ bathymetry dataset. Model values were interpolated to the observational dataset's longitudes and latitudes using 3D interpolation. **Note:** this analysis has been restricted to observations on the shelf region. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 563 values extracted from the observational database. The map below shows the locations of the matched up data for chlorophyll concentration.

The following model output was used to compare with observational values: **P1_Chl+P2_Chl+P3_Chl+P4_Chl**.

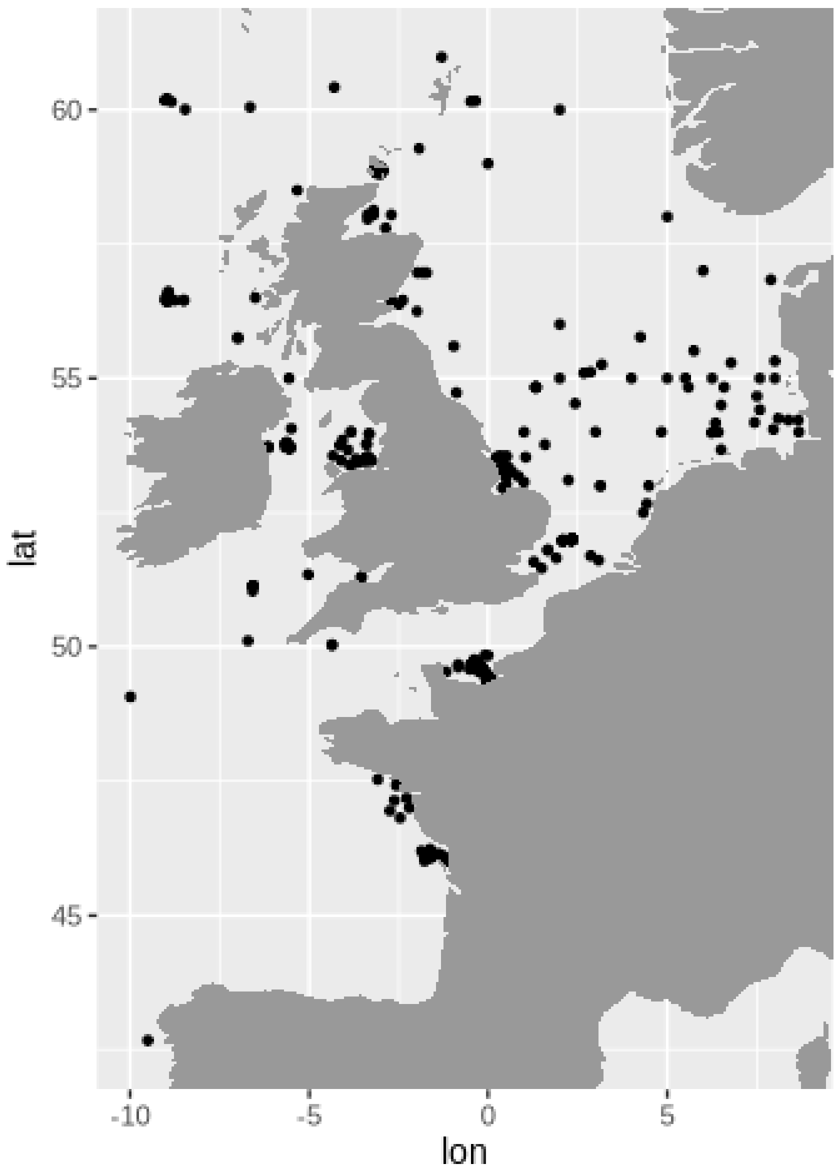


Figure 9.4: Locations of matchups between simulated and observed chlorophyll concentration near the bottom of the

water column.

The number of observations in each month ranged from 15 in November to 99 in May. Figure 9.5 below shows the distribution of observations in each month.

Figure 9.5 below shows the bias between the model and observational data for chlorophyll concentration. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

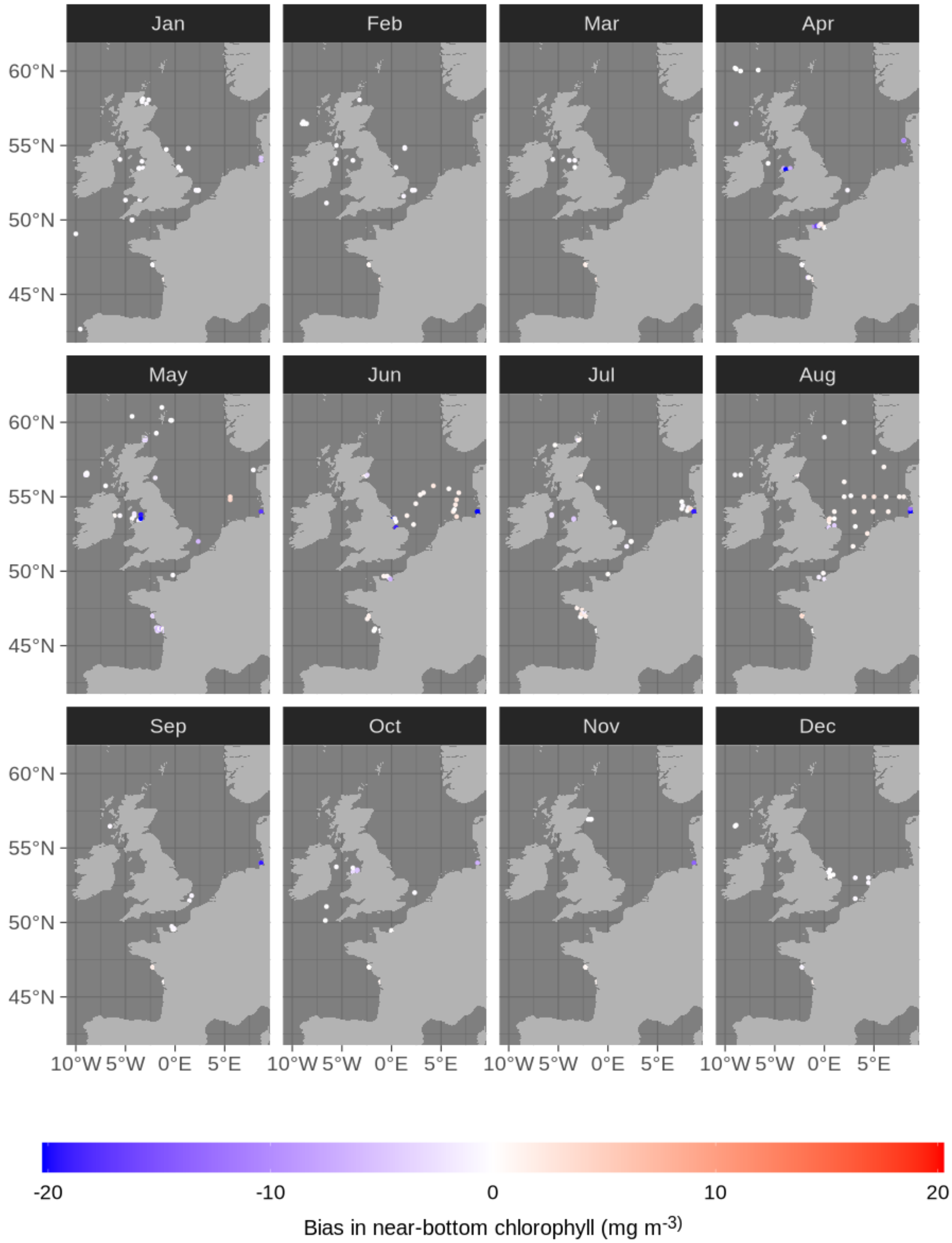


Figure 9.5: Bias in near-bottom chlorophyll concentration. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 9.6 and 9.7 show the distribution of near-bottom chlorophyll concentration observations in the model and observational datasets. This is shown for each month of the year (Figure 9.6) and for the entire year (Figure 9.7).

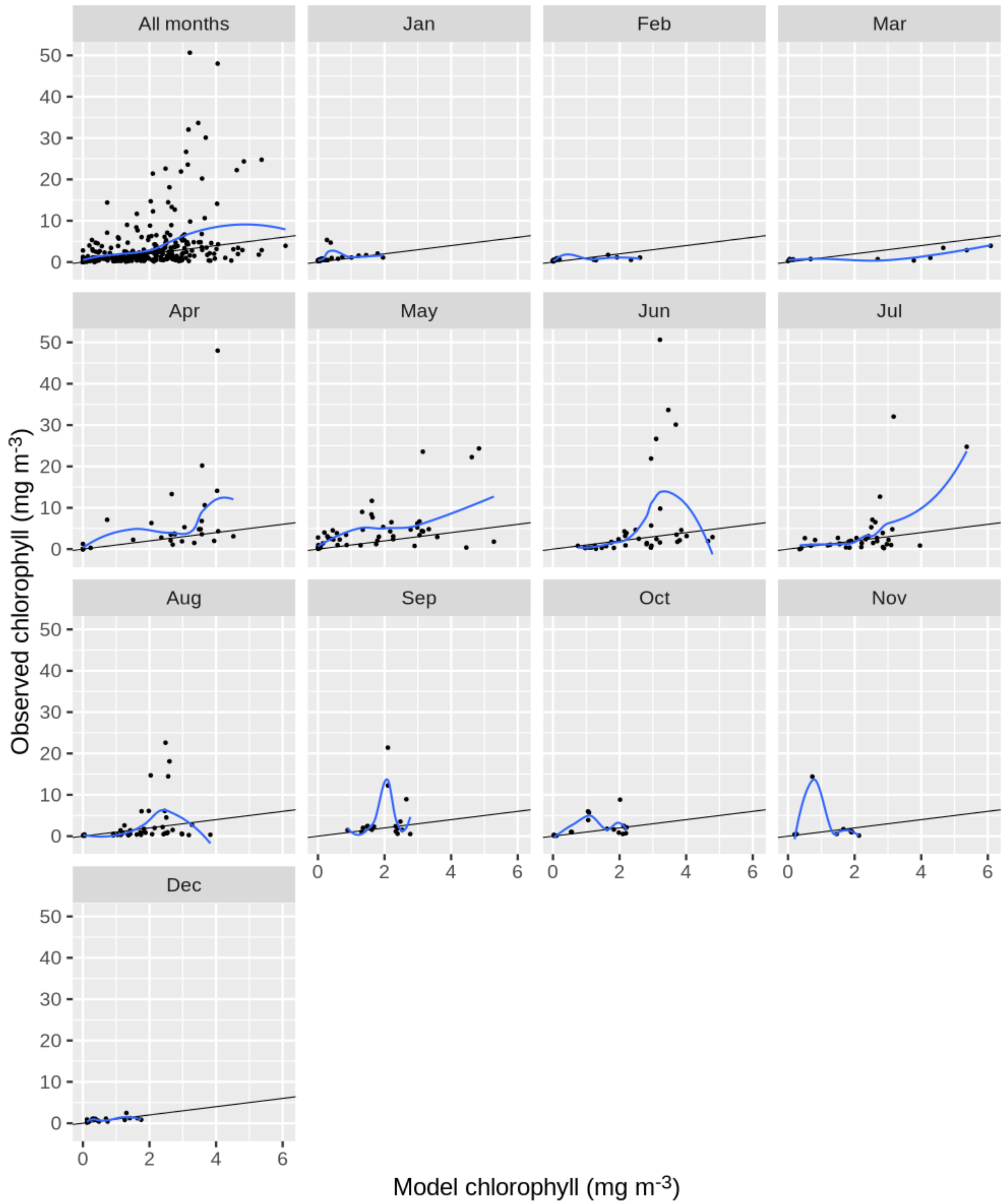


Figure 9.6: Simulated versus observed chlorophyll concentration near the bottom of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

9.4 Summary statistics for near-bottom chlorophyll

The overall ability of the model to predict the observed chlorophyll concentration was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	-1.83	7.22	0.35	563
Jan	-0.49	1.08	0.40	51
Feb	0.07	0.87	0.46	37
Mar	1.71	2.37	0.77	18
Apr	-2.30	7.19	0.46	53
May	-3.32	6.57	0.48	99
Jun	-4.62	16.08	0.19	57
Jul	-1.96	7.40	0.54	64
Aug	-1.52	6.29	0.32	72
Sep	-1.09	4.55	0.08	27
Oct	-1.11	2.44	0.17	49
Nov	-0.33	3.65	-0.21	15
Dec	-0.14	0.76	0.43	21

Table 9.3: Average bias (mg m^{-3}) and root-mean square deviation (mg m^{-3}) for the model's near-bottom chlorophyll concentration for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed chlorophyll concentration was performed. The modelled chlorophyll concentration was used as the independent variable and the observed chlorophyll concentration was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	1.91	0.14	0.16	0.00
Jan	0.72	0.68	0.11	0.04
Feb	0.27	0.42	0.38	0.00
Mar	0.45	0.24	0.60	0.01
Apr	2.54	-0.36	0.16	0.03
May	1.92	1.20	0.28	0.00
Jun	3.33	-2.55	0.10	0.05
Jul	3.23	-3.43	0.26	0.00
Aug	1.52	0.07	0.08	0.07
Sep	1.46	1.05	0.02	0.57
Oct	0.66	1.37	0.05	0.35
Nov	-1.00	3.34	0.04	0.58
Dec	0.47	0.55	0.25	0.04

Table 9.4: Linear regression analysis of modelled and observed chlorophyll. The modelled chlorophyll concentration was used as the independent variable and the observed chlorophyll concentration was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

9.5 Depth-resolved comparisons of modelled and observed chlorophyll

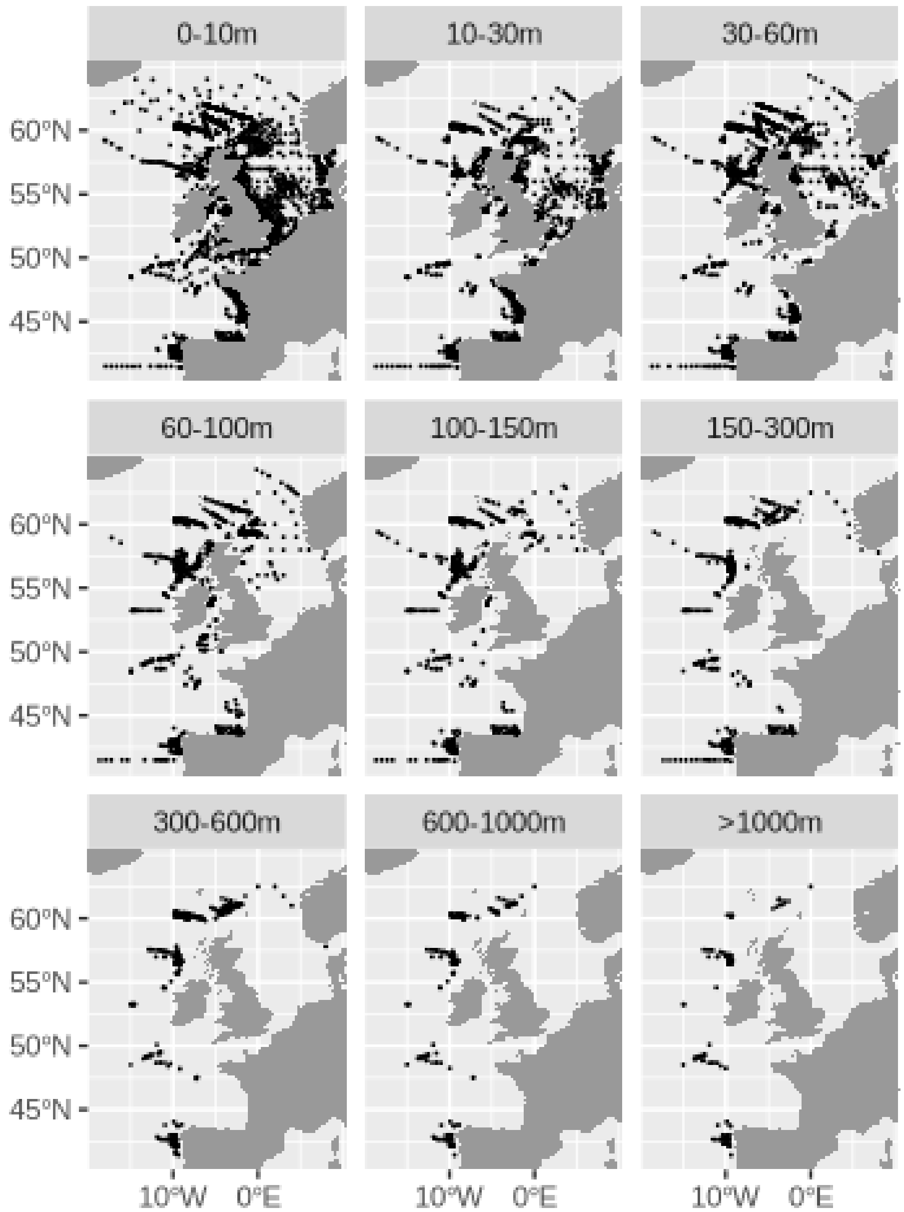


Figure 9.7: The geographic distribution of matchups between the model and observational chlorophyll concentration. The data has been binned into depth ranges. The depth ranges are 0-10m, 10-30m, 30-60m, 60-100m, 100-150m, 150-300m, 300-600m, 600-1000m, and >1000m. The number of observations in each depth range is shown in the tables below.

Depth	Bias	*r*	RMSD	Number of observations
0-150m	0.04	0.30	1.92	29,442
0-10m	-0.07	0.26	2.78	7,698
10-30m	0.15	0.19	2.31	6,599
30-60m	0.28	0.07	1.34	6,376
60-100m	-0.09	0.02	0.55	5,126
100-150m	-0.18	-0.15	0.42	3,643
150-300m	-0.19	-0.08	0.35	3,481
300-600m	-0.21	-0.03	0.41	2,618
600-1000m	-0.26	-0.09	0.47	1,232
>1000m	-0.21	-0.05	0.48	476

Table 9.5: Average bias (mg m^{-3}), root-mean square difference (RMSD) and correlation coefficient of modelled and observed chlorophyll concentration for different depth ranges. The bias is calculated as model-observation. The RMSD is the square root of the mean squared difference. The correlation coefficient is the Pearson correlation coefficient between the model and observed values.

9.6 Data Sources for validation of chlorophyll

ICES Data Portal, Dataset on Ocean HydroChemistry, Extracted March 3, 2023. ICES, Copenhagen

CHAPTER TEN

SEA SURFACE CHLOROPHYLL VALIDATION USING GRIDDED OBSERVATIONS FROM NSBC

We used version 1.1 of the **North Sea Biogeochemical Climatology** (NSBC) to validate **sea surface chlorophyll**. NSBC is a monthly climatology that covers the region 47°-65°N and 15°W-15°E. The data is made up of observations over the period 1960-2014. For validation purposes we only used the level 3 data, which a gridded monthly climatology at a spatial resolution of 1/4°. The data can be download from NSBC.

Matchup procedure: The model and observations were matched up as follows. First, the model dataset was cropped by a small amount to make sure cells close to the boundary were removed. The model was then regridded to the observational grid if the observational grid was coarser using nearest neighbour. Only grid cells with model and observational data were maintained. The following model output was used to compare with the observational values: **P1_Chl+P2_Chl+P3_Chl+P4_Chl**.

10.1 Baseline climatologies of sea surface chlorophyll

Climatologies of model and observational sea surface chlorophyll concentration are shown in the figures below. The model climatology is calculated using the years **1986-2017**.

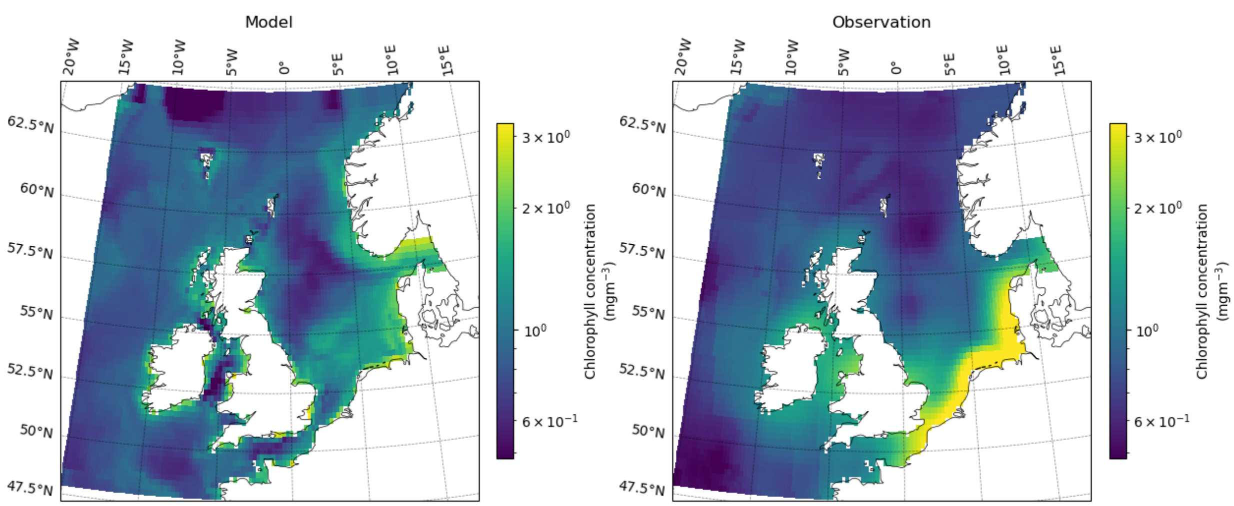


Figure 10.1: Annual average surface chlorophyll concentration from the model (1986-2017) and observations. Data is limited to the 2nd and 98th percentile of the combined model and observational data. Arrows indicate that values can exceed the colourbar limits.

10.2 Assessing model bias for surface chlorophyll concentration

Figure 10.2 shows the average bias of surface chlorophyll concentration simulated by the model. A positive bias indicates that the model overestimates the observation, while a negative bias indicates that the model overpredicts the observation.

The spatial average bias of surface chlorophyll concentration is -0.04 mgm^{-3}. Overall, the model overestimates the observations in 59.8% of the model domain.

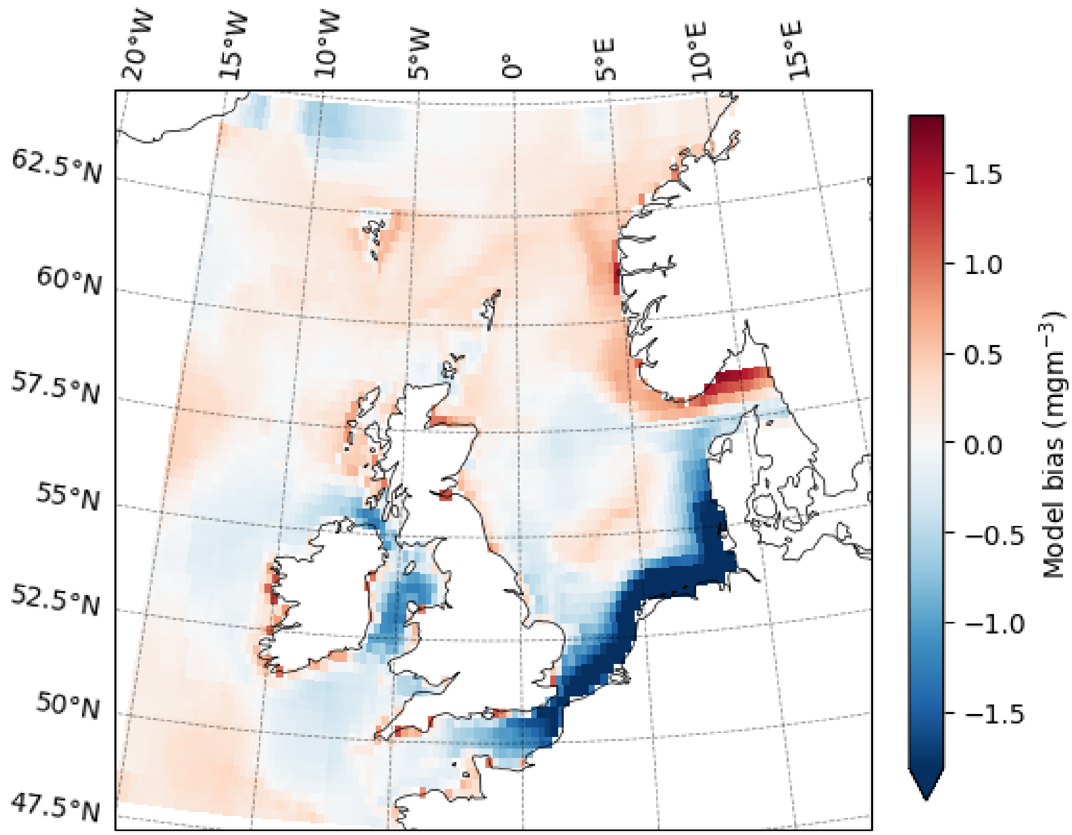


Figure 10.2: Bias of surface chlorophyll concentration from the model. A positive bias indicates that the model overestimates the observation. For clarity, the colourbar is limited to the 2nd and 98th percentile of the data.

10.3 Can the model reproduce seasonality of sea surface chlorophyll concentration?

The ability of the model to reproduce seasonality of sea surface chlorophyll concentration was assessed by comparing the modelled and observed seasonal cycle of chlorophyll concentration. First, we derive a monthly climatology for the model data. Then, we calculate the Pearson correlation coefficient between the modelled and observed chlorophyll concentration at each grid cell.

Note: we are only assessing the ability of the model to reproduce the ability of the model to reproduce seasonal changes, not long-term trends.

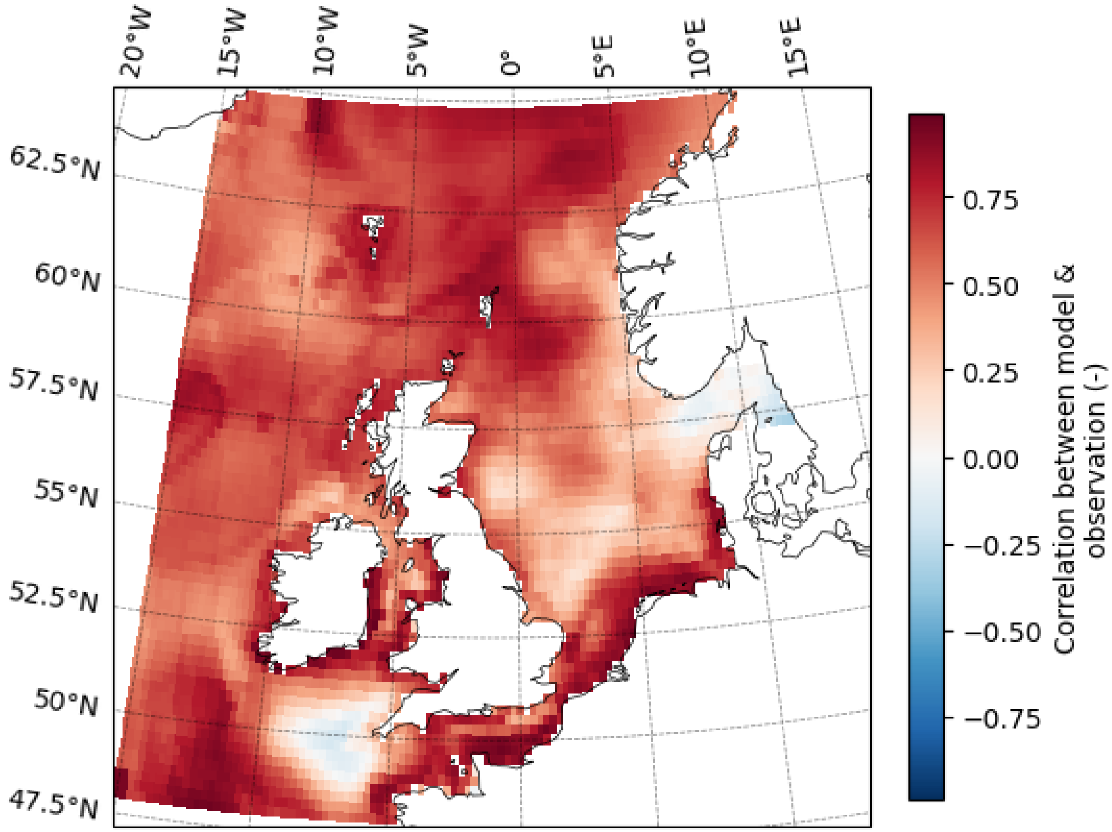
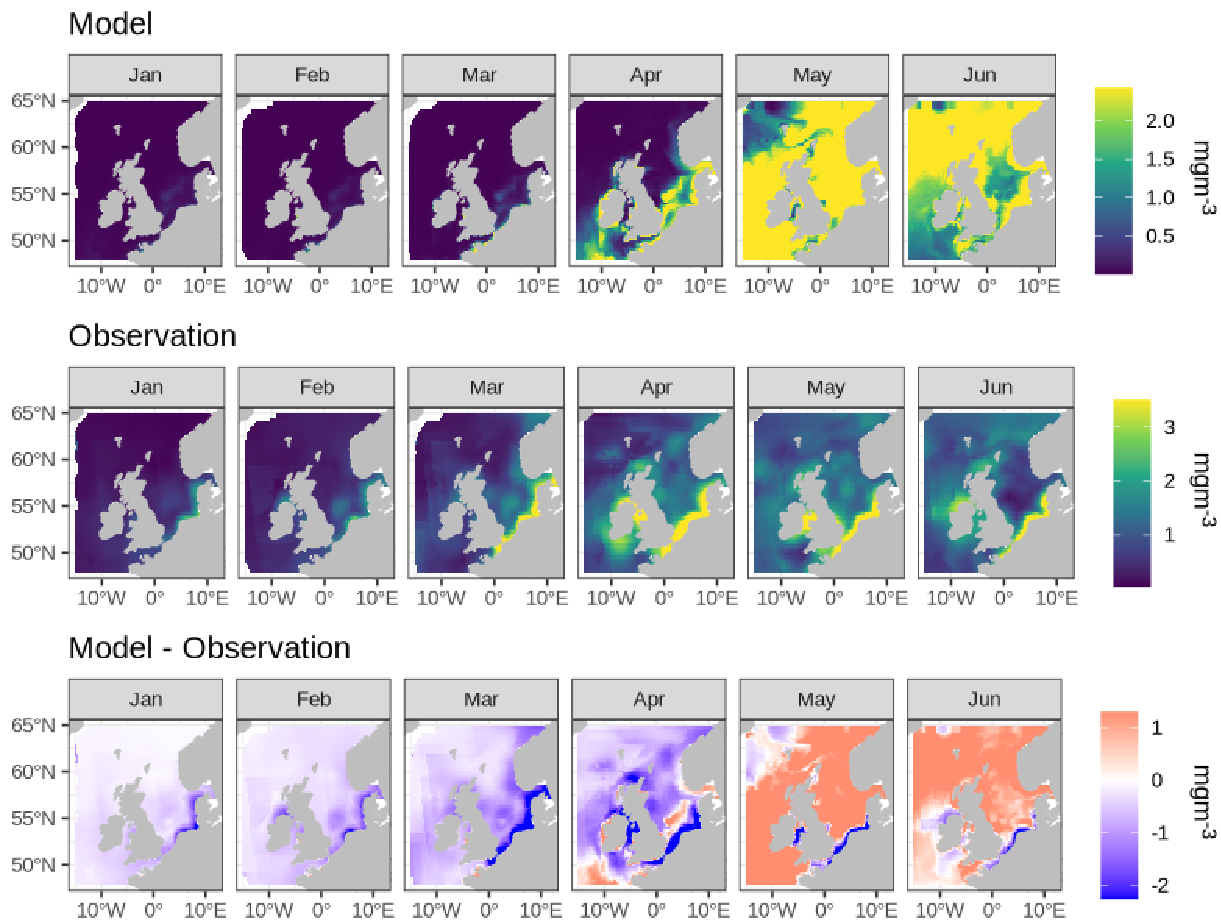


Figure 10.3: Seasonal temporal correlation between model and observations for surface chlorophyll concentration. This is the Pearson correlation coefficient between climatology monthly mean values in the model and observations.

The seasonal cycles of simulated and observed chlorophyll concentration are compared in Figure 10.4 below. This figure shows the model and observation average in each month of the year, and the differences between the two each month


Model
Jan
Feb
Mar
Apr
May
Jun
65°N
60°N
55°N
50°N
10°W 0° 10°E
2.0
1.5
1.0
0.5
mgm-3
Observation
3
2
1
Model - Observation
1
0
-1
-2

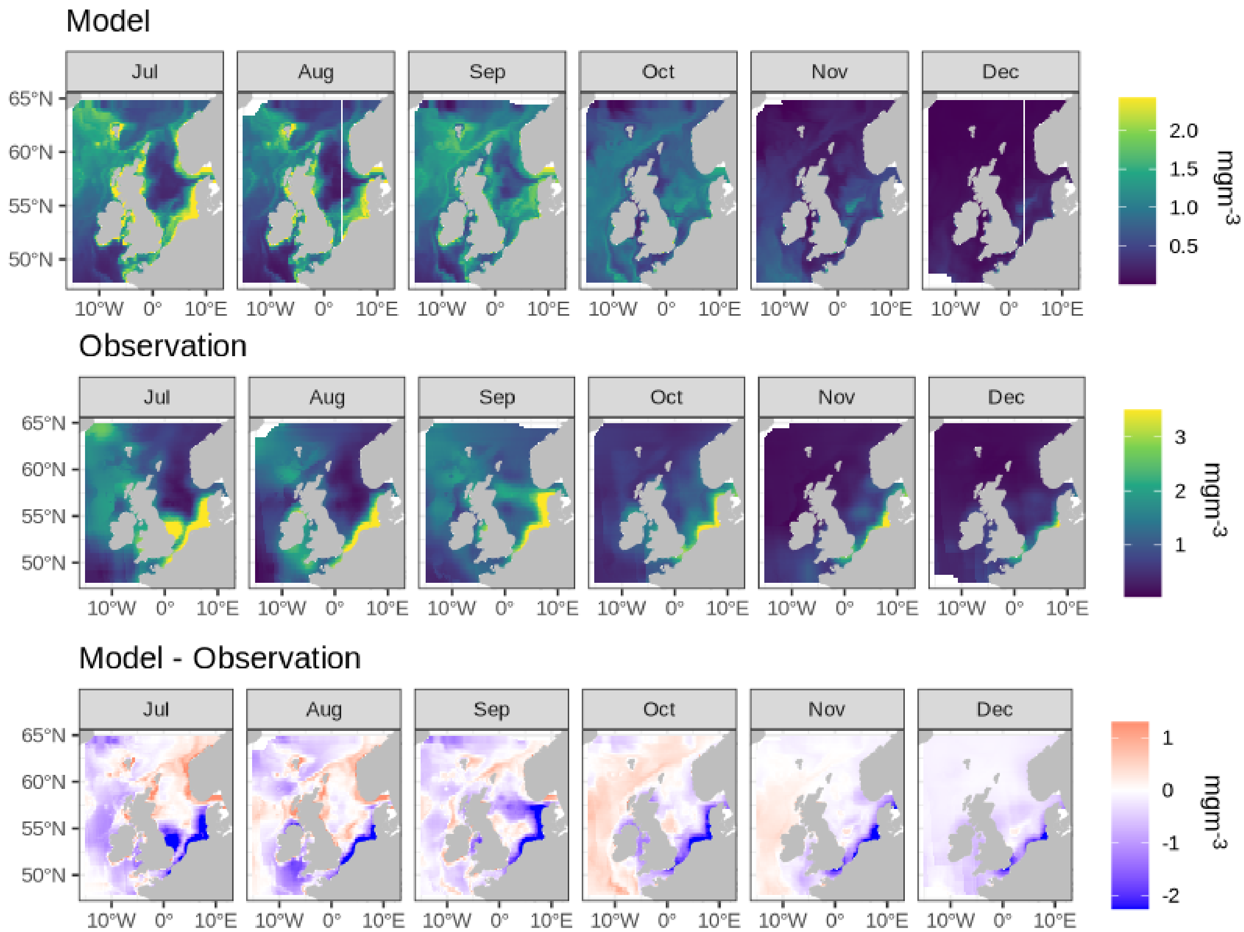


Figure 10.4: Monthly mean surface chlorophyll concentration for the model, observation and the difference between model and observations. For clarity, the maximum values are capped to the 98th percentiles.

10.4 Regional assessment of model performance for sea surface chlorophyll concentration

We assessed the regional performance of the model by comparing the model with observations in a number of regions. The regions considered are mapped below.

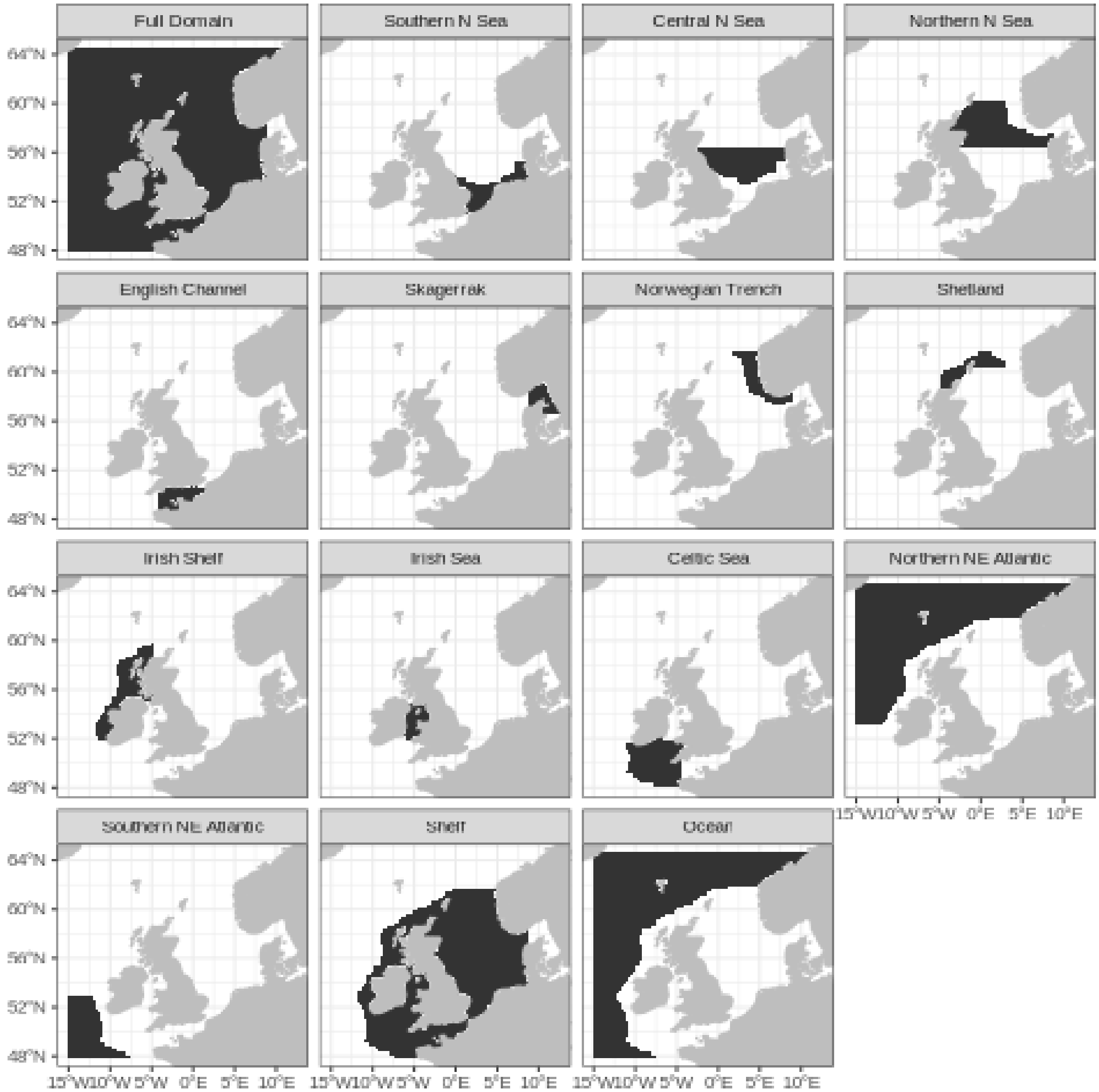


Figure 10.5: Regions used for validation of sea surface chlorophyll concentration.

Time series were constructed comparing the monthly mean of the spatial average sea surface chlorophyll concentration in each region. The spatial average was calculated using the mean of all grid cells within each region, accounting for grid cell area.

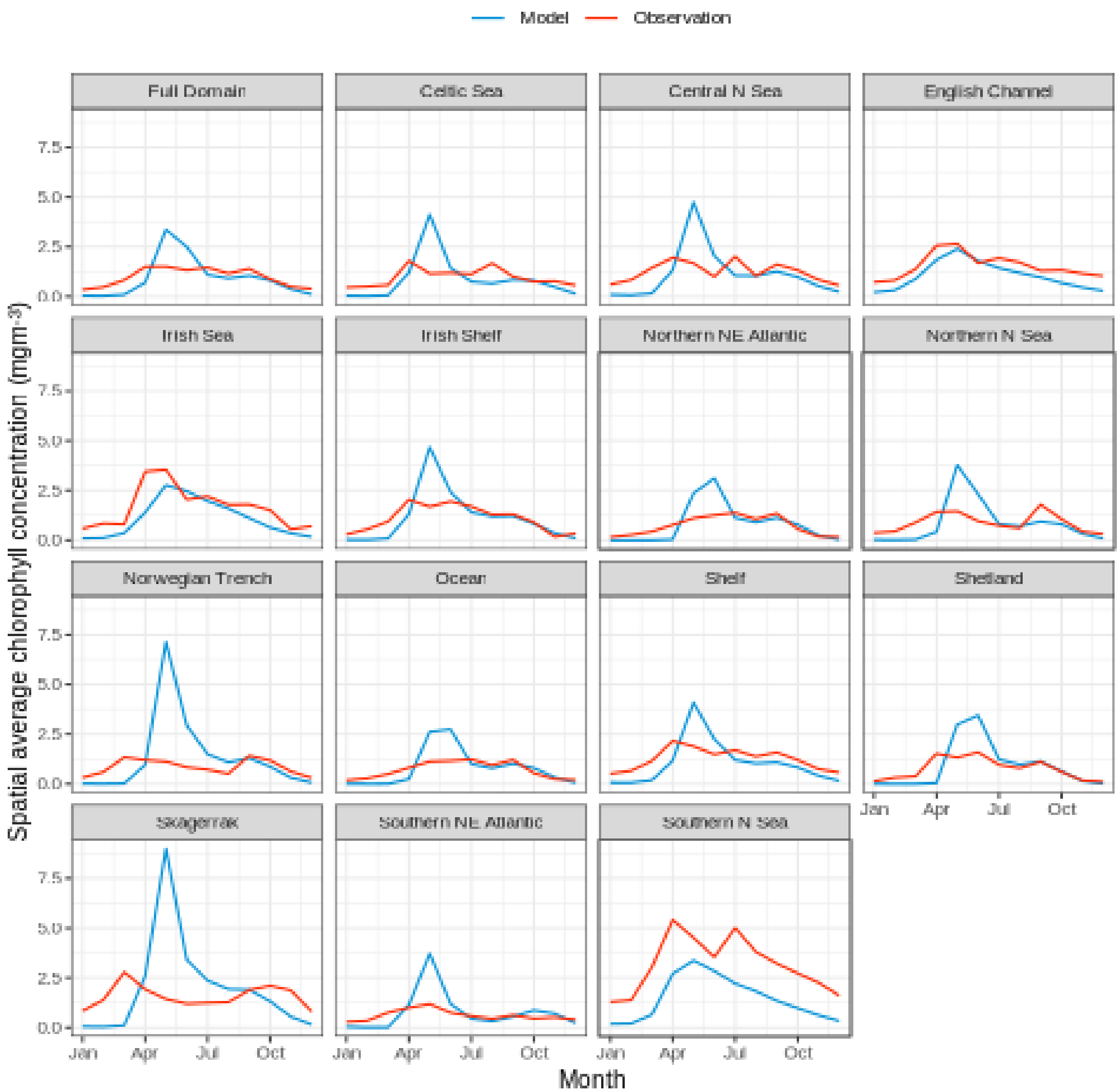


Figure 10.6: Seasonal cycle of sea surface chlorophyll concentration for model and observations for each region. The spatial average is taken over the region.

The table below shows the average bias of sea surface chlorophyll concentration in each region. The bias is calculated as the modelled value minus the observed value. A positive bias indicates that the model overestimates the observed value,

while a negative bias indicates that the model underestimates the observed value.

Region	Temporal correlation	Bias	RMSD
Full Domain	0.58	-0.05	1.15
Celtic Sea	0.38	-0.08	1.09
Central North Sea	0.38	-0.11	1.34
Channel	0.73	-0.49	1.00
Irish Sea	0.67	-0.57	1.19
Irish Shelf	0.61	0.04	1.21
Northern North East Atlantic	0.64	0.09	0.90
Northern North Sea	0.52	-0.02	1.09
Norwegian Trench	0.30	0.50	2.00
Ocean	0.65	0.10	0.89
Shelf	0.52	-0.21	1.37
Shetland	0.69	0.13	0.92
Skagerrak	-0.04	0.39	2.71
Southern North East Atlantic	0.70	0.16	0.86
Southern North Sea	0.69	-1.71	2.41

Table 10.1: Summary of performance of the model sea surface chlorophyll concentration in each region. The bias (mgm-3) column represents the spatial average of the annual mean modelled value minus the observed value. The temporal correlation column represents the spatial mean of the temporal correlation between the model and observations per grid cell.

10.5 Can the model reproduce spatial patterns of sea surface chlorophyll concentration?

The ability of the model to reproduce spatial patterns of sea surface chlorophyll concentration was assessed by comparing the modelled and observed chlorophyll concentration at each grid cell. We calculated the Pearson correlation coefficient between the modelled and observed chlorophyll concentration at each grid cell.

This was carried out monthly and using the annual mean in each grid cell

Time period	r
Annual mean	0.59
Jan	0.55
Feb	0.44
Mar	0.47
Apr	0.57
May	0.09
Jun	0.18
Jul	0.60
Aug	0.56
Sep	0.41
Oct	0.30
Nov	0.48
Dec	0.60

Table 10.2: Pearson correlation coefficient between modelled and observed sea surface chlorophyll concentration at each grid cell. The correlation was calculated monthly and using the annual mean in each grid cell.

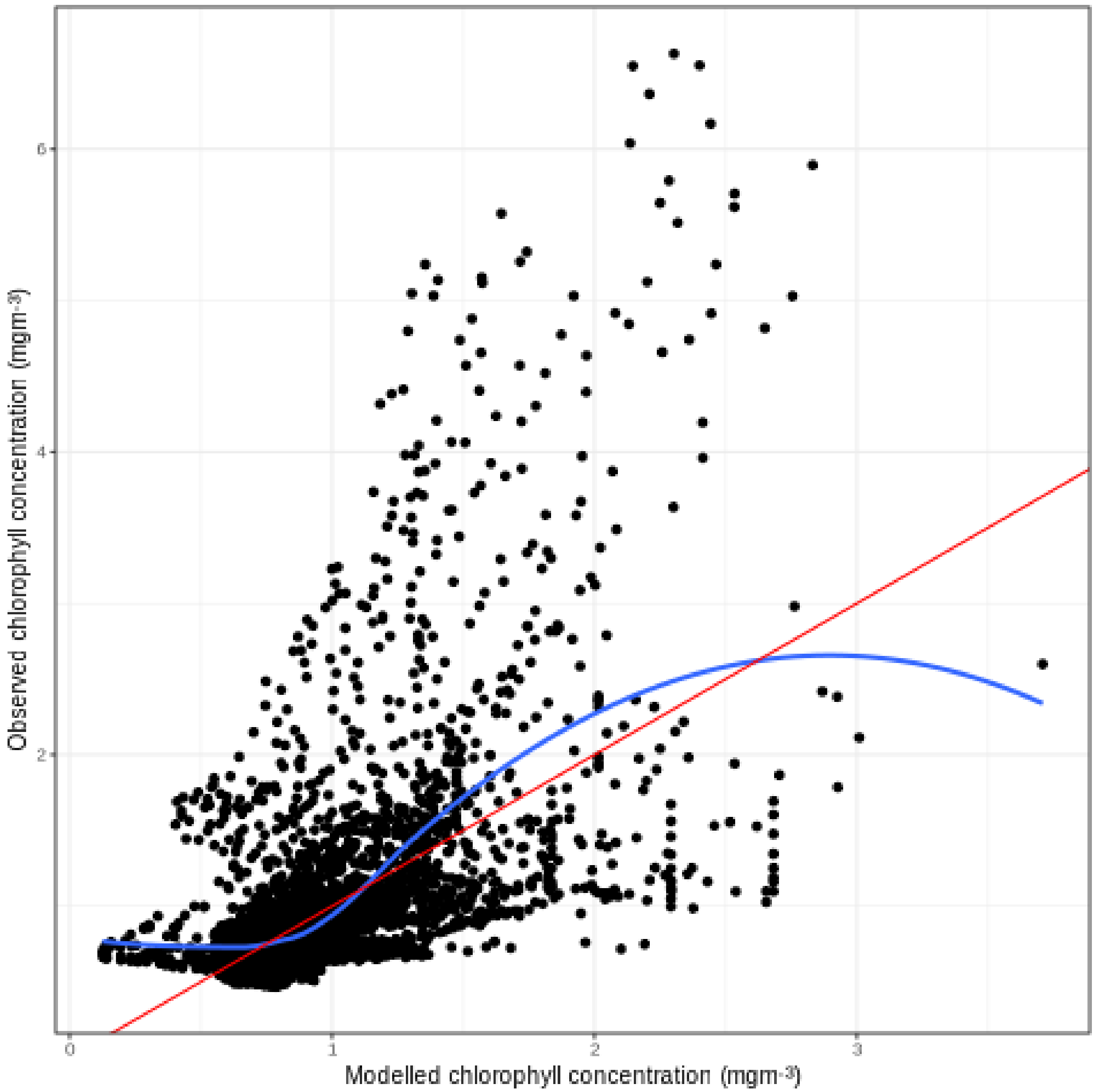


Figure 10.7: Scatter plot of modelled and observed annual average sea surface chlorophyll concentration in grid cells. The red line is the 1:1 line. The blue line is the linear regression of the modelled and observed values. The slope of the blue line is the slope of the regression line.

10.6 Data Sources for validation of chlorophyll concentration

Hinrichs, Iris; Gouretski, Viktor; Paetsch, Johannes; Emeis, Kay; Stammer, Detlef (2017). North Sea Biogeochemical Climatology (Version 1.1).

URL: https://www.cen.uni-hamburg.de/en/icdc/data/ocean/nsbc.html

CHAPTER

ELEVEN

VALIDATION OF DOC USING POINT OBSERVATIONS

11.1 Performance of model sea surface DOC

This data was extracted from vertical profiles. Values from the **top 5m** were extracted from the database. This was compared with the model values from the sea surface level. Dissolved organic carbon data was compiled from multiple sources The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 4,813 values extracted from the observational database. The map below shows the locations of the matched up data for DOC.

The following model output was used to compare with observational values: **R1_c+R^2_c+R3_c**.

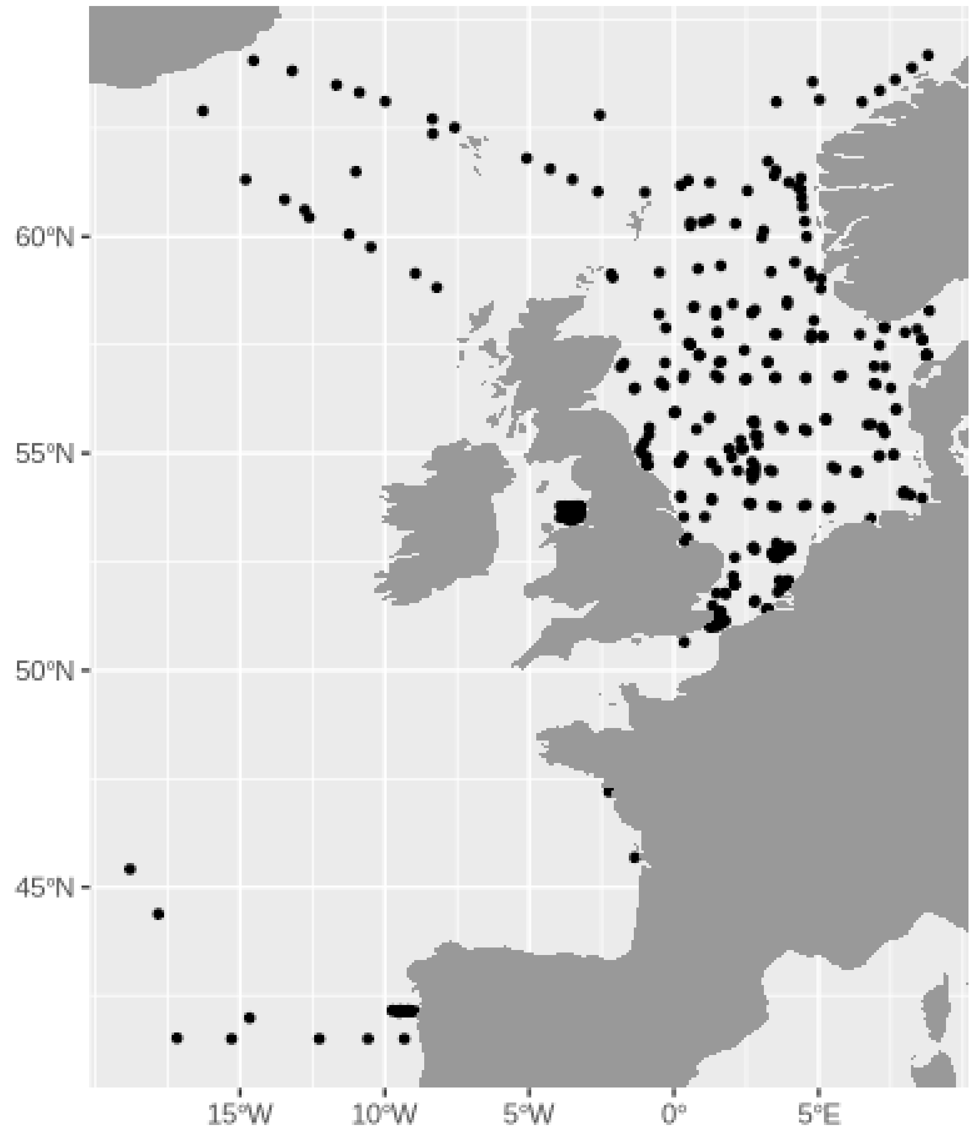


Figure 11.1: Locations of matchups between simulated and observed DOC in the top 5m of the water column.

The number of observations in each month ranged from 279 in February to 610 in August. Figure 11.2 below shows the distribution of observations in each month.

Figure 11.2 below shows the bias between the model and observational data for DOC. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

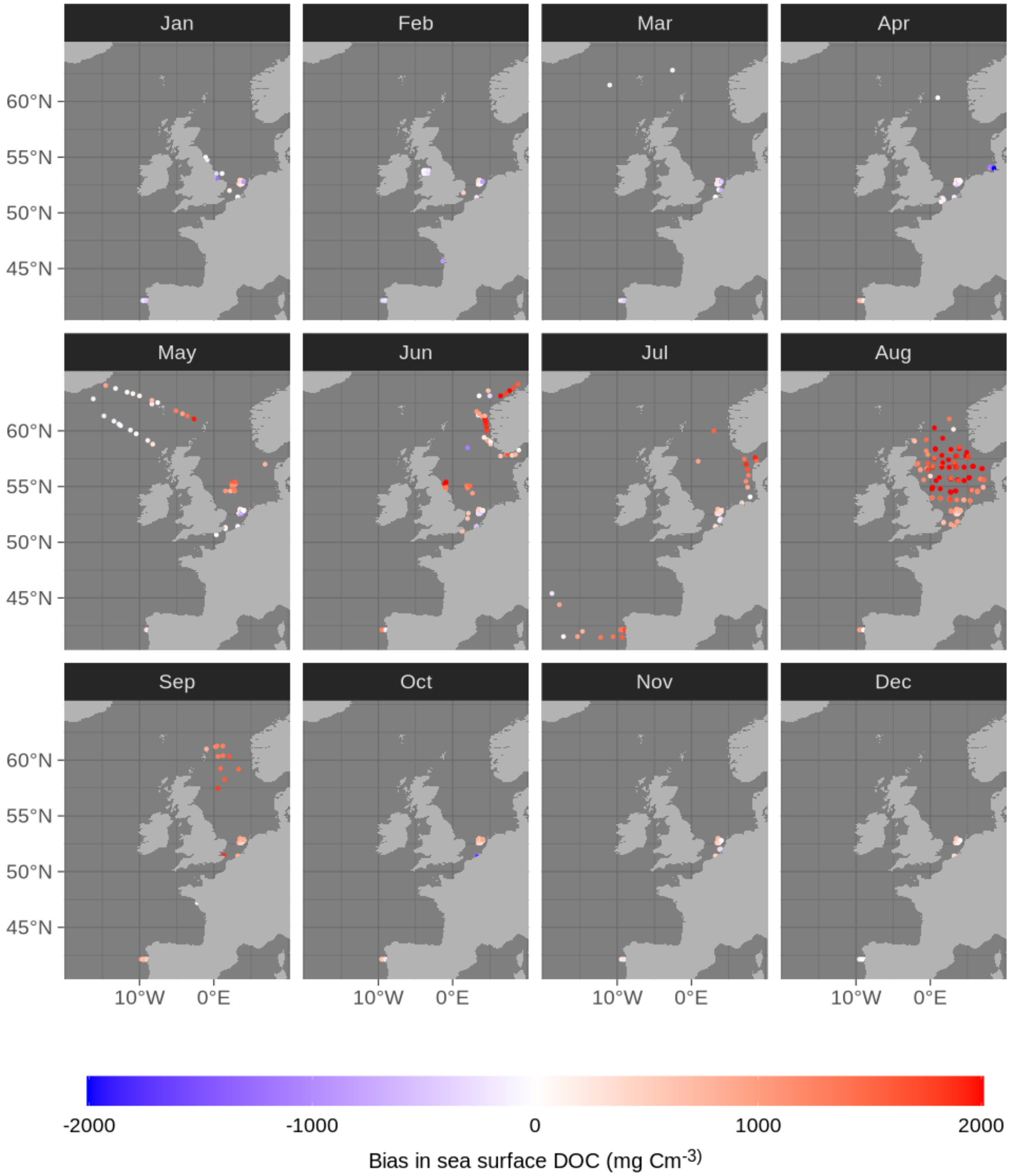


Figure 11.2: Bias in sea surface DOC. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 11.3 and 11.4 show the distribution of sea surface DOC observations in the model and observational datasets. This is shown for each month of the year (Figure 11.3) and for the entire year (Figure 11.4).

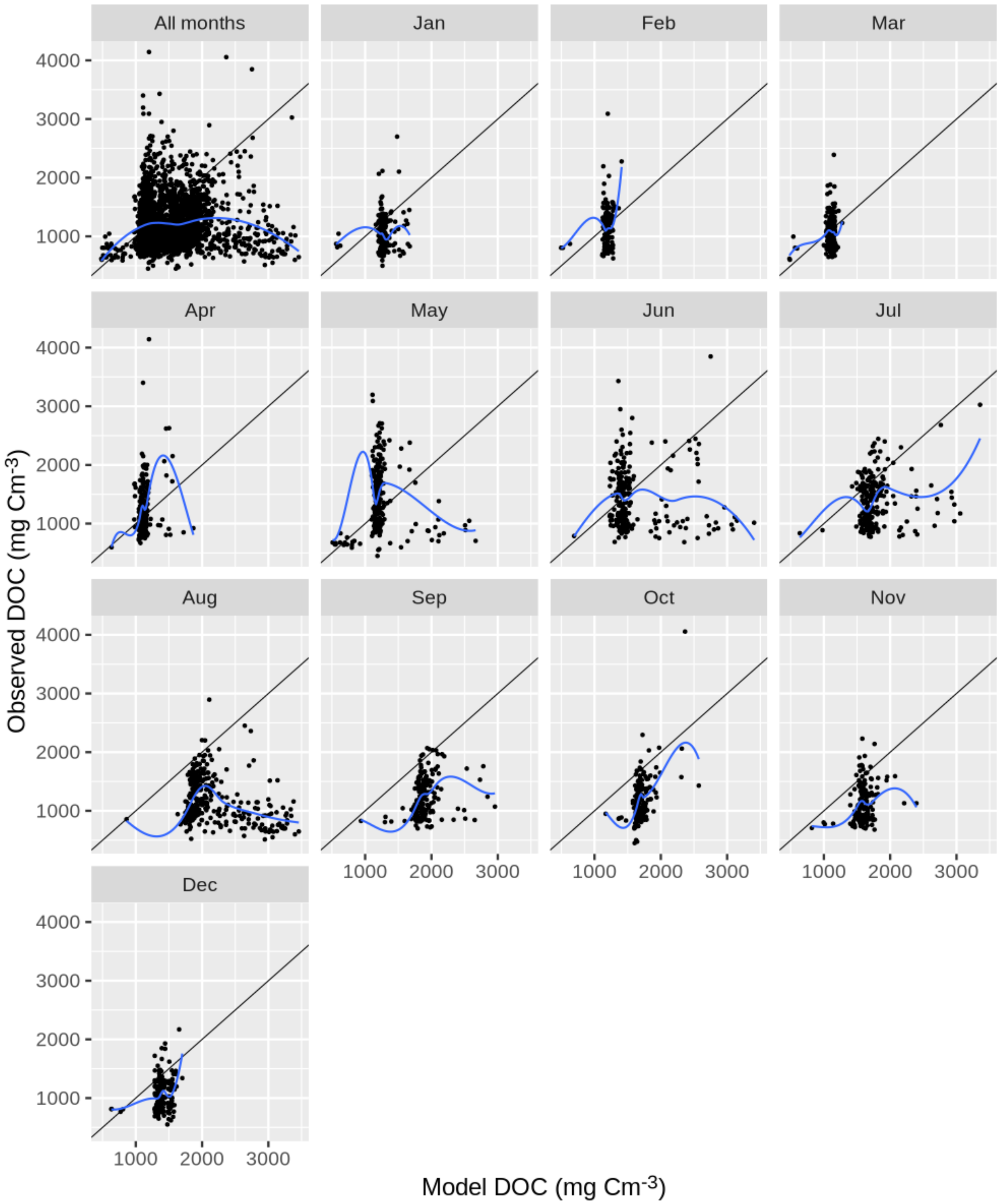


Figure 11.3: Simulated versus observed DOC in the top 5m of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

11.2 Summary statistics for sea surface DOC

The overall ability of the model to predict the observed DOC was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	232.45	616.11	0.04	4,813
Jan	166.11	365.08	0.07	403
Feb	11.23	346.75	0.12	279
Mar	-16.92	311.66	0.06	383
Apr	-166.82	494.78	0.05	387
May	-296.41	664.22	0.01	494
Jun	60.23	642.20	-0.06	455
Jul	382.12	621.03	0.09	480
Aug	789.91	996.02	-0.22	610
Sep	599.98	704.88	0.23	327
Oct	432.62	557.72	0.44	333
Nov	415.88	513.55	0.20	328
Dec	289.91	421.38	0.20	334

Table 11.1: Average bias (mg Cm^{-3}) and root-mean square deviation (mg Cm^{-3}) for the model's sea surface DOC for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed DOC was performed. The modelled DOC was used as the independent variable and the observed DOC was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.04	1154.45	0.00	0.04
Jan	0.24	718.79	0.01	0.13
Feb	0.61	417.12	0.03	0.03
Mar	0.44	586.30	0.03	0.03
Apr	0.61	563.38	0.03	0.02
May	0.01	1382.55	0.00	0.92
Jun	-0.04	1491.51	0.00	0.63
Jul	0.25	908.24	0.04	0.00
Aug	-0.23	1667.99	0.07	0.00
Sep	0.34	583.84	0.05	0.00
Oct	1.37	-1141.90	0.27	0.00
Nov	0.49	349.87	0.06	0.00
Dec	0.43	448.35	0.05	0.01

Table 11.2: Linear regression analysis of modelled and observed DOC. The modelled DOC was used as the independent variable and the observed DOC was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

11.3 Performance of model near-bottom doc

CHAPTER

TWELVE

VALIDATION OF MESOZOOPLANKTON BIOMASS USING POINT OBSERVATIONS

12.1 Performance of model sea surface mesozooplankton concentration

This data was extracted from vertical profiles. Values from the **top 5m** were extracted from the database. This was compared with the model values from the sea surface level. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 54,415 values extracted from the observational database. The map below shows the locations of the matched up data for mesozooplankton concentration.

The following model output was used to compare with observational values: **Z4_c**.

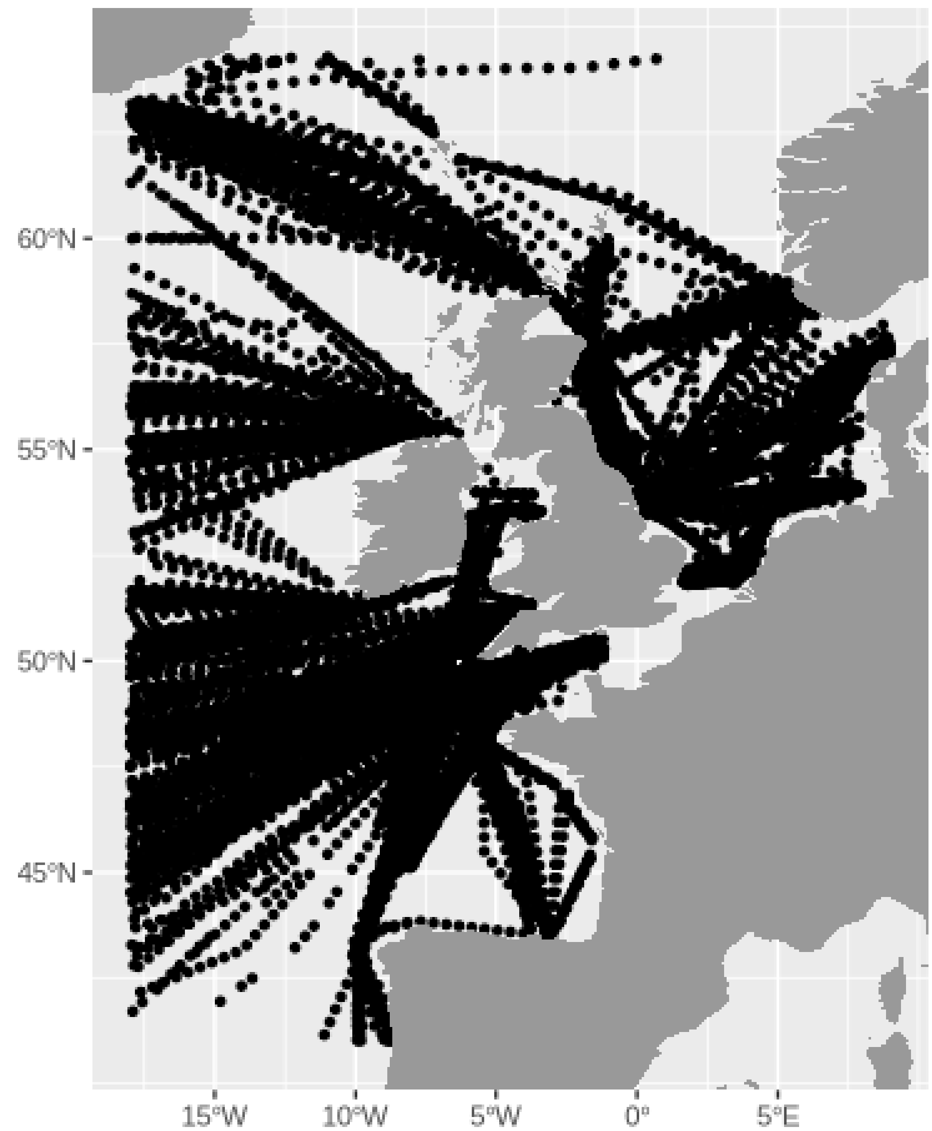


Figure 12.1: Locations of matchups between simulated and observed mesozooplankton concentration in the top 5m of the water column.

The number of observations in each month ranged from 3,402 in January to 5,082 in September. Figure 12.2 below shows the distribution of observations in each month.

Figure 12.2 below shows the bias between the model and observational data for mesozooplankton concentration. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

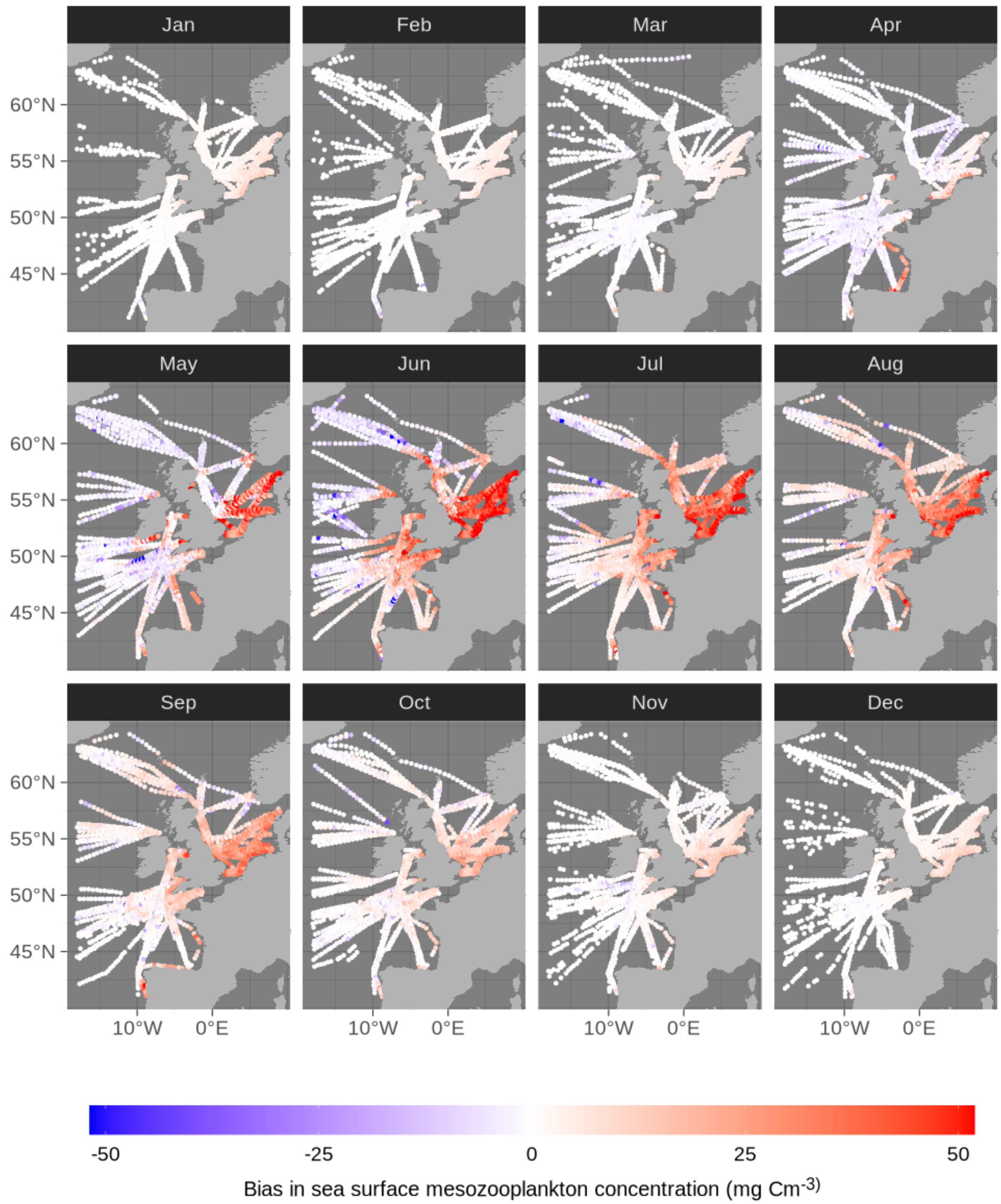


Figure 12.2: Bias in sea surface mesozooplankton concentration. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 12.3 and 12.4 show the distribution of sea surface mesozooplankton concentration observations in the model and observational datasets. This is shown for each month of the year (Figure 12.3) and for the entire year (Figure 12.4).

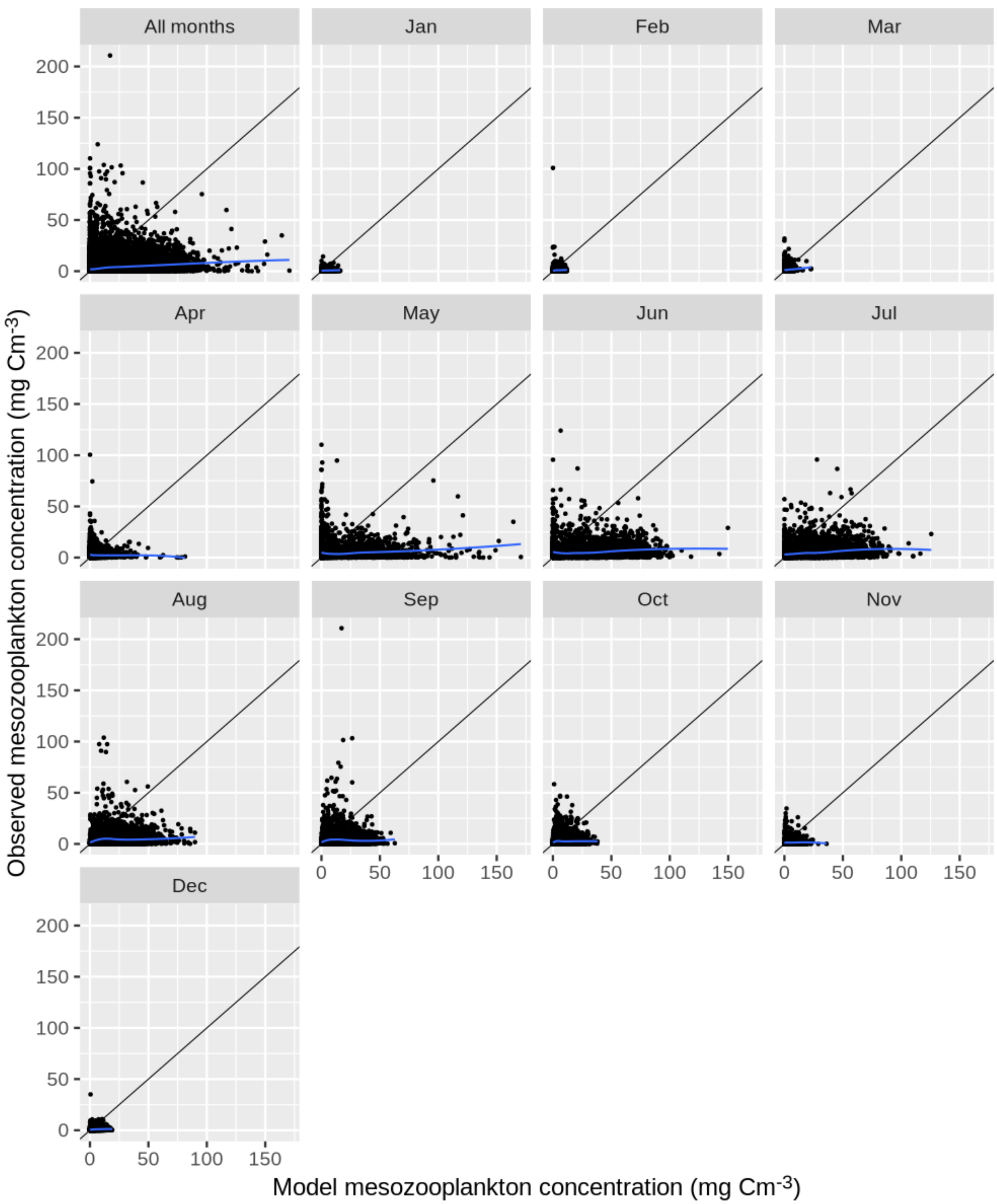


Figure 12.3: Simulated versus observed mesozooplankton concentration in the top 5m of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

12.2 Summary statistics for sea surface mesozooplankton concentration

The overall ability of the model to predict the observed mesozooplankton concentration was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	7.34	16.50	0.22	54,415
Jan	1.98	3.53	0.20	3,402
Feb	1.21	3.19	0.10	3,985
Mar	0.30	2.52	0.16	4,417
Apr	-0.21	6.71	0.01	4,816
May	8.11	23.00	0.05	4,808
Jun	18.19	29.43	0.09	4,666
Jul	18.01	26.00	0.21	4,917
Aug	15.56	22.62	0.08	4,861
Sep	10.61	16.96	0.02	5,082
Oct	5.07	9.36	0.05	4,899
Nov	2.69	5.51	0.08	4,366
Dec	2.13	4.09	0.17	4,196

Table 12.1: Average bias (mg Cm^{-3}) and root-mean square deviation (mg Cm^{-3}) for the model's sea surface mesozooplankton concentration for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed mesozooplankton concentration was performed. The modelled mesozooplankton concentration was used as the independent variable and the observed mesozooplankton concentration was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.08	1.93	0.05	0.00
Jan	0.05	0.39	0.04	0.00
Feb	0.08	0.55	0.01	0.00
Mar	0.13	0.94	0.02	0.00
Apr	0.01	2.48	0.00	0.28
May	0.02	4.33	0.00	0.00
Jun	0.03	4.46	0.01	0.00
Jul	0.07	2.93	0.04	0.00
Aug	0.03	3.41	0.01	0.00
Sep	0.01	3.31	0.00	0.22
Oct	0.03	2.17	0.00	0.00
Nov	0.04	1.08	0.01	0.00
Dec	0.06	0.56	0.03	0.00

Table 12.2: Linear regression analysis of modelled and observed mesozooplankton concentration. The modelled mesozooplankton concentration was used as the independent variable and the observed mesozooplankton concentration was

used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

12.3 Performance of model near-bottom mesozoo

CHAPTER

THIRTEEN

VALIDATION OF NITRATE USING POINT OBSERVATIONS

13.1 Performance of model sea surface nitrate

Values from the **top 5m** of the water column were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 20,506 values extracted from the observational database. The map below shows the locations of the matched up data for nitrate concentration.

The following model output was used to compare with observational values: **N3_n**.

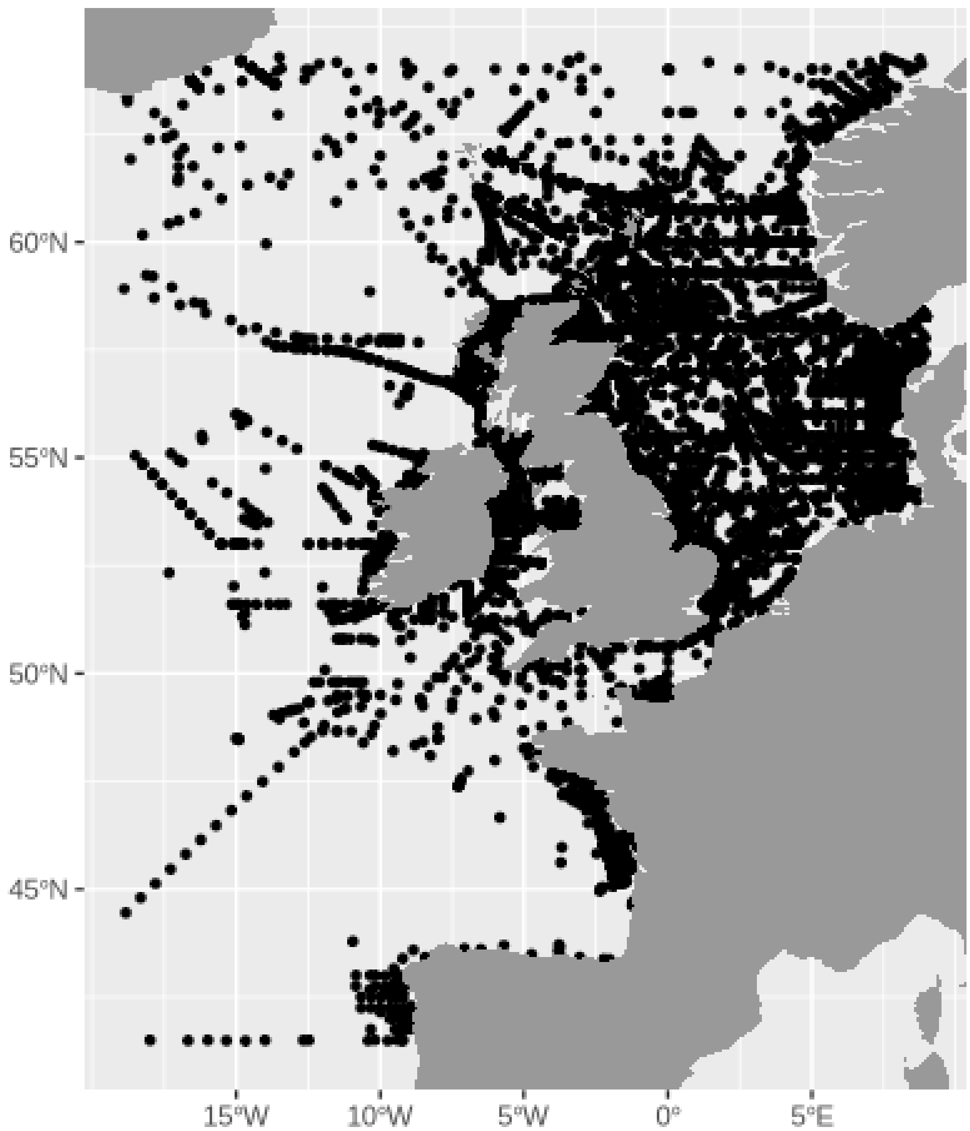


Figure 13.1: Locations of matchups between simulated and observed nitrate concentration in the top 5m of the water column.

The number of observations in each month ranged from 845 in March to 2,672 in May. Figure 13.2 below shows the distribution of observations in each month.

Figure 13.2 below shows the bias between the model and observational data for nitrate concentration. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

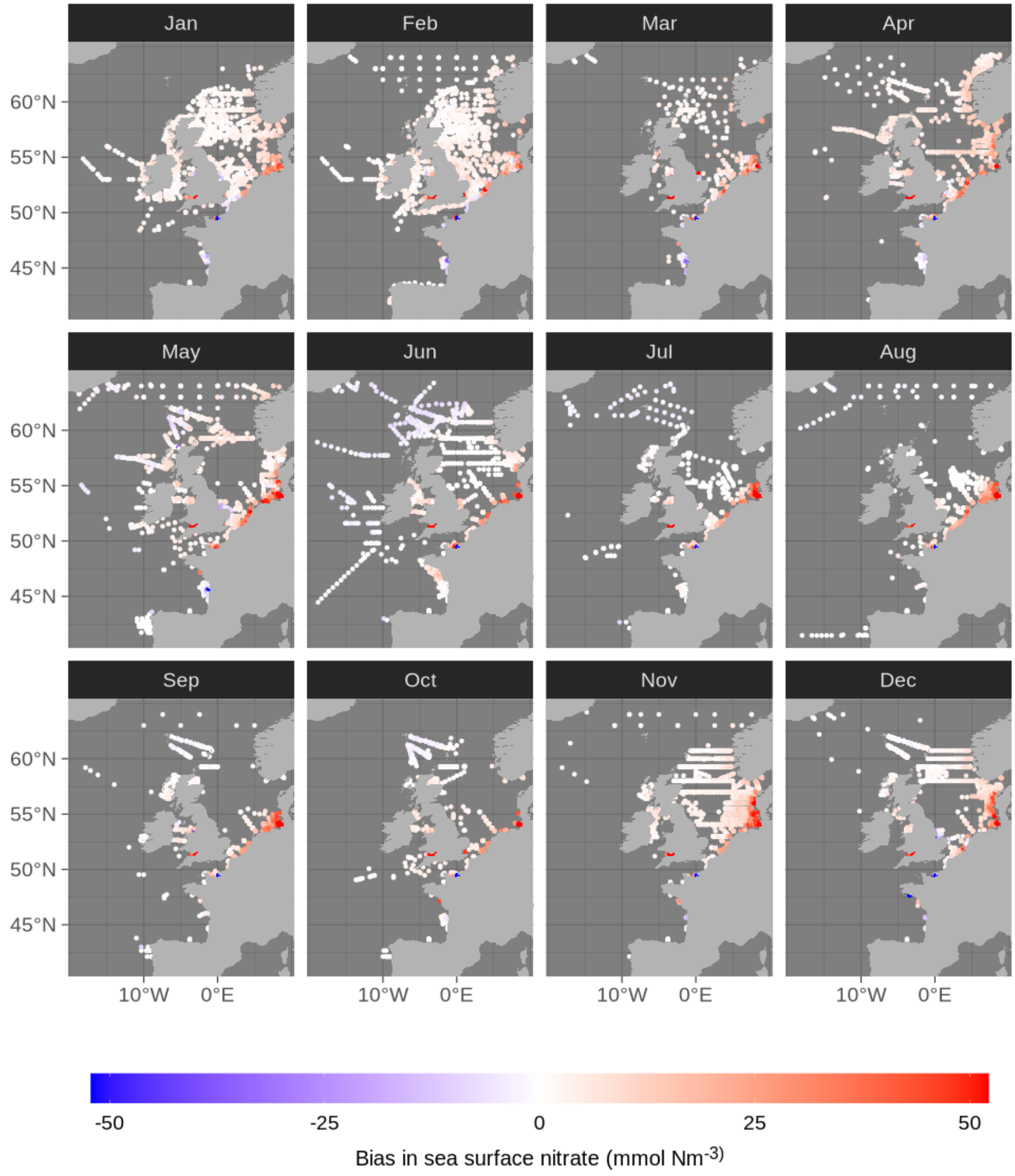


Figure 13.2: Bias in sea surface nitrate concentration. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 13.3 and 13.4 show the distribution of sea surface nitrate concentration observations in the model and observational datasets. This is shown for each month of the year (Figure 13.3) and for the entire year (Figure 13.4).

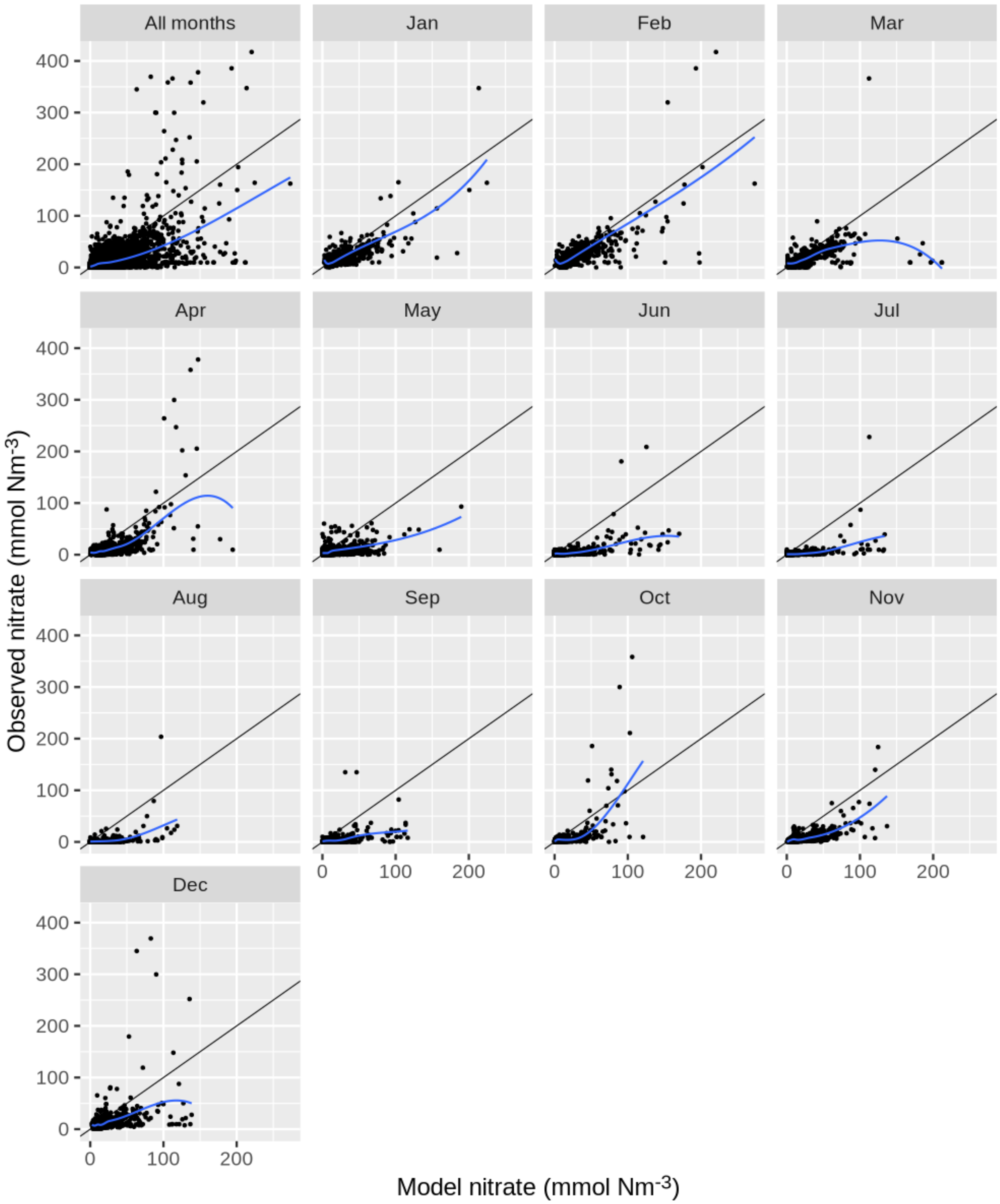


Figure 13.3: Simulated versus observed nitrate concentration in the top 5m of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

13.2 Summary statistics for sea surface nitrate

The overall ability of the model to predict the observed nitrate concentration was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	7.16	17.90	0.54	20,506
Jan	3.64	11.87	0.74	2,649
Feb	4.05	15.71	0.73	2,260
Mar	8.39	26.84	0.48	845
Apr	9.81	19.69	0.62	1,931
May	6.21	14.60	0.49	2,672
Jun	8.19	19.72	0.51	1,985
Jul	10.28	21.55	0.44	1,148
Aug	11.01	21.22	0.42	1,047
Sep	9.68	19.40	0.40	946
Oct	4.97	17.29	0.54	1,117
Nov	10.05	16.26	0.57	2,191
Dec	5.36	19.75	0.41	1,715

Table 13.1: Average bias (mmol Nm^{-3}) and root-mean square deviation (mmol Nm^{-3}) for the model's sea surface nitrate concentration for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed nitrate concentration was performed. The modelled nitrate concentration was used as the independent variable and the observed nitrate concentration was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.48	1.03	0.33	0.00
Jan	0.65	1.20	0.62	0.00
Feb	0.72	0.73	0.59	0.00
Mar	0.33	7.78	0.22	0.00
Apr	0.73	-4.35	0.40	0.00
May	0.22	3.01	0.21	0.00
Jun	0.24	0.65	0.27	0.00
Jul	0.21	-0.26	0.19	0.00
Aug	0.21	-0.60	0.18	0.00
Sep	0.18	0.97	0.15	0.00
Oct	0.74	-1.56	0.31	0.00
Nov	0.33	0.77	0.38	0.00
Dec	0.49	3.32	0.19	0.00

Table 13.2: Linear regression analysis of modelled and observed nitrate. The modelled nitrate concentration was used as the independent variable and the observed nitrate concentration was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

13.3 Performance of model near-bottom nitrate

Near-bottom values of nitrate were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The near-bottom was defined as observations **within 2m of the seabed**. This was interpolated to the observational grid using the $GEBCO_2$ bathymetry dataset. Model values were interpolated to the observational dataset's longitudes and latitudes using 3D interpolation. **Note:** this analysis has been restricted to observations on the shelf region. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 1,908 values extracted from the observational database. The map below shows the locations of the matched up data for nitrate concentration.

The following model output was used to compare with observational values: **N3_n**.

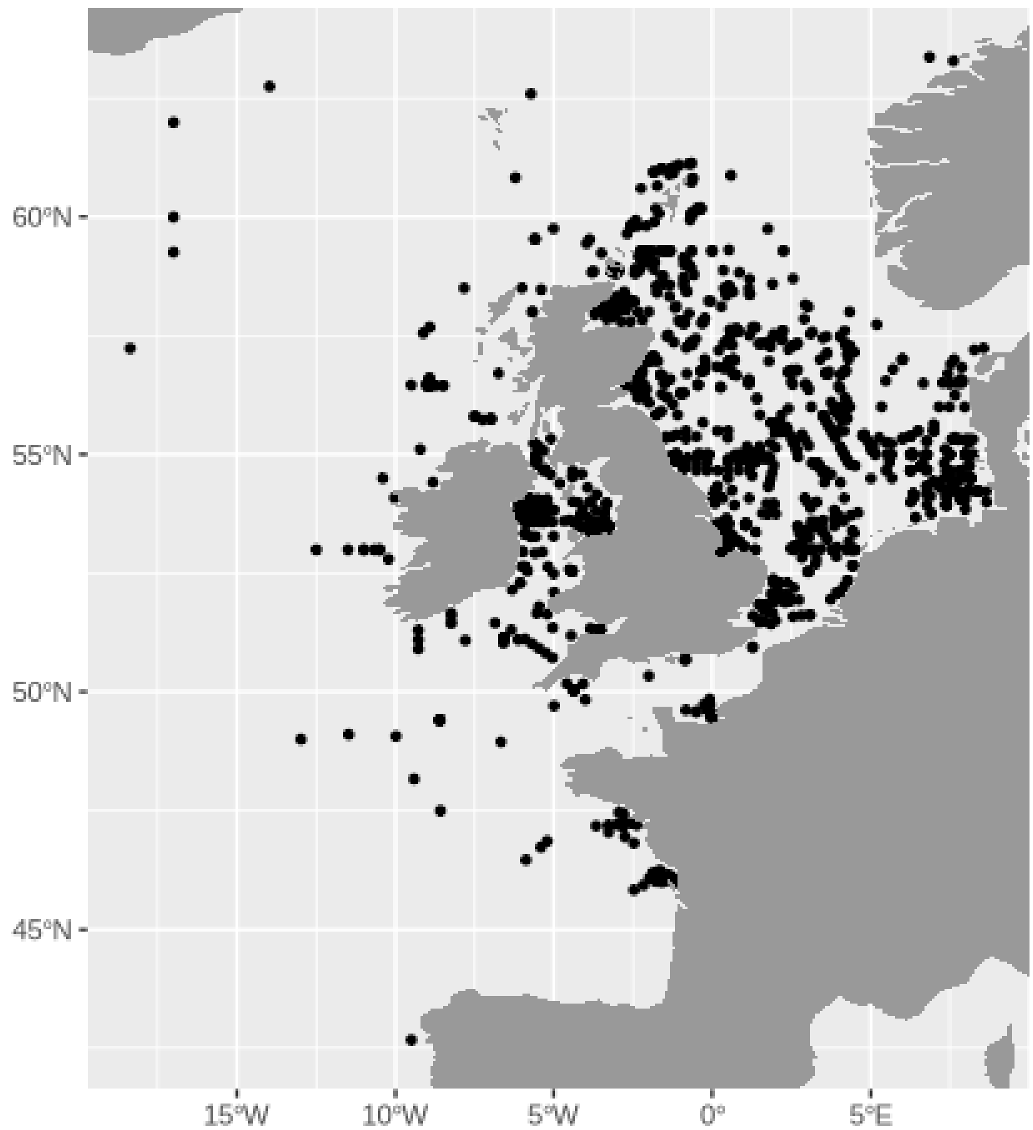


Figure 13.4: Locations of matchups between simulated and observed nitrate concentration near the bottom of the water

column.

The number of observations in each month ranged from 49 in December to 371 in February. Figure 13.5 below shows the distribution of observations in each month.

Figure 13.5 below shows the bias between the model and observational data for nitrate concentration. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

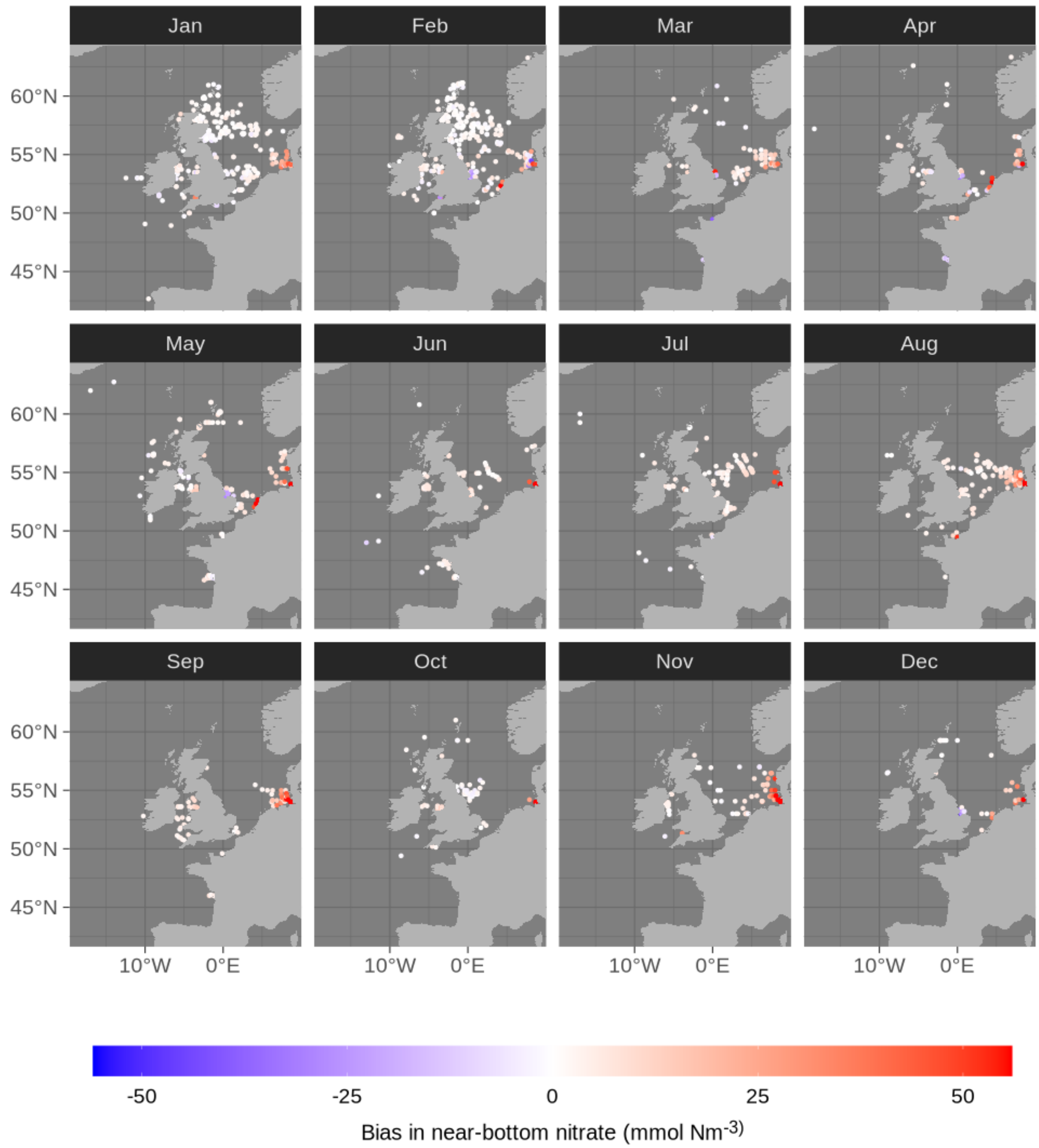


Figure 13.5: Bias in near-bottom nitrate concentration. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 13.6 and 13.7 show the distribution of near-bottom nitrate concentration observations in the model and observational datasets. This is shown for each month of the year (Figure 13.6) and for the entire year (Figure 13.7).

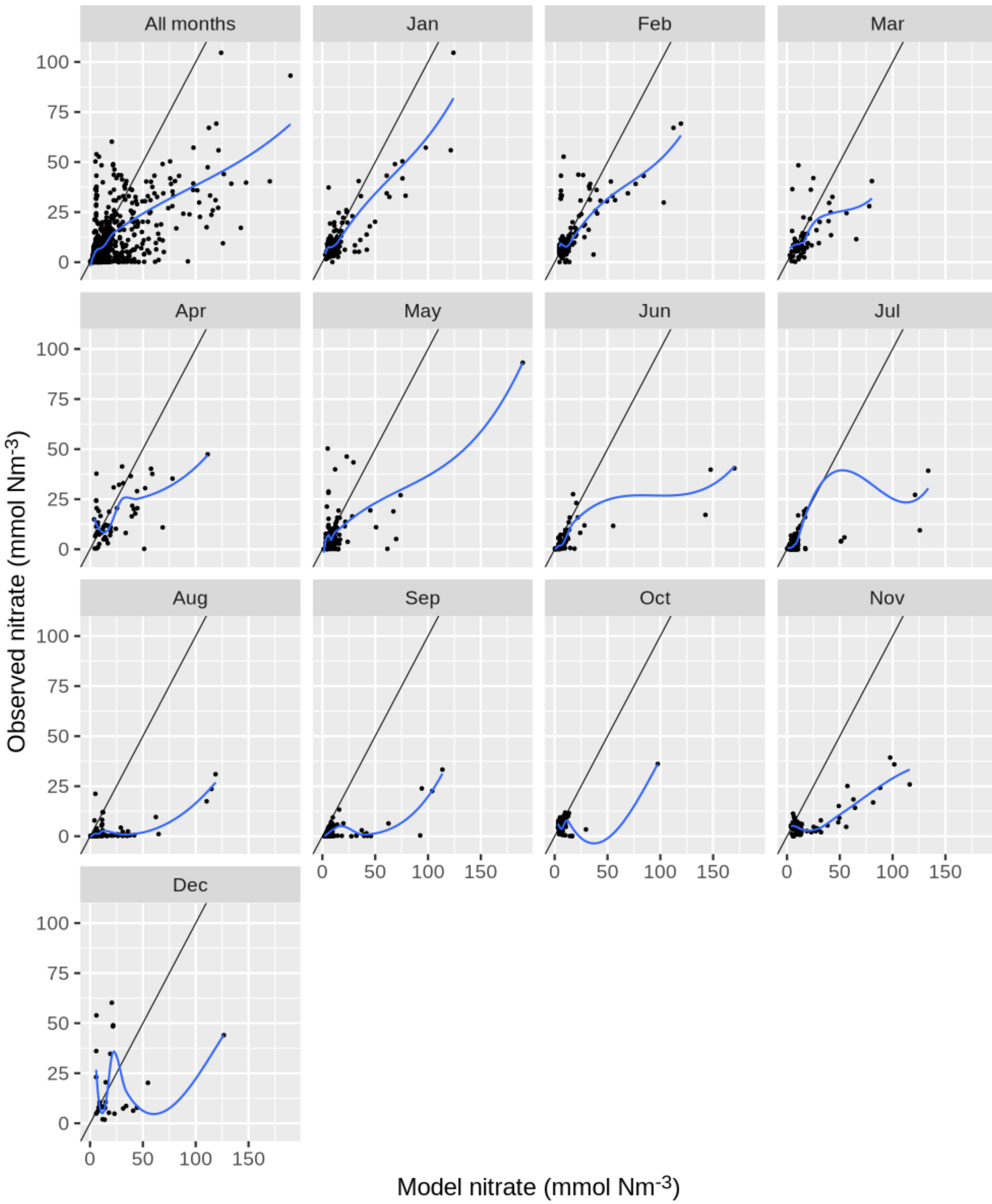


Figure 13.6: Simulated versus observed nitrate concentration near the bottom of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

13.4 Summary statistics for near-bottom nitrate

The overall ability of the model to predict the observed nitrate concentration was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	9.03	20.10	0.57	1,908
Jan	4.05	11.09	0.81	360
Feb	2.90	10.34	0.76	371
Mar	9.08	16.19	0.60	89
Apr	7.22	17.96	0.53	101
May	7.70	18.57	0.68	223
Jun	19.35	37.55	0.79	102
Jul	12.98	26.01	0.62	174
Aug	17.97	28.89	0.73	140
Sep	16.83	25.60	0.55	105
Oct	8.58	14.44	0.36	99
Nov	16.76	24.54	0.79	95
Dec	10.01	24.41	0.16	49

Table 13.3: Average bias (mmol Nm^{-3}) and root-mean square deviation (mmol Nm^{-3}) for the model's near-bottom nitrate concentration for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed nitrate concentration was performed. The modelled nitrate concentration was used as the independent variable and the observed nitrate concentration was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.31	4.01	0.37	0.00
Jan	0.55	2.96	0.75	0.00
Feb	0.50	4.49	0.56	0.00
Mar	0.35	6.16	0.28	0.00
Apr	0.33	8.32	0.35	0.00
May	0.38	3.33	0.40	0.00
Jun	0.21	2.48	0.62	0.00
Jul	0.19	1.07	0.46	0.00
Aug	0.16	-0.21	0.46	0.00
Sep	0.16	0.40	0.45	0.00
Oct	0.30	2.99	0.55	0.00
Nov	0.24	1.74	0.66	0.00
Dec	0.20	12.62	0.07	0.13

Table 13.4: Linear regression analysis of modelled and observed nitrate. The modelled nitrate concentration was used as the independent variable and the observed nitrate concentration was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

13.5 Depth-resolved comparisons of modelled and observed nitrate

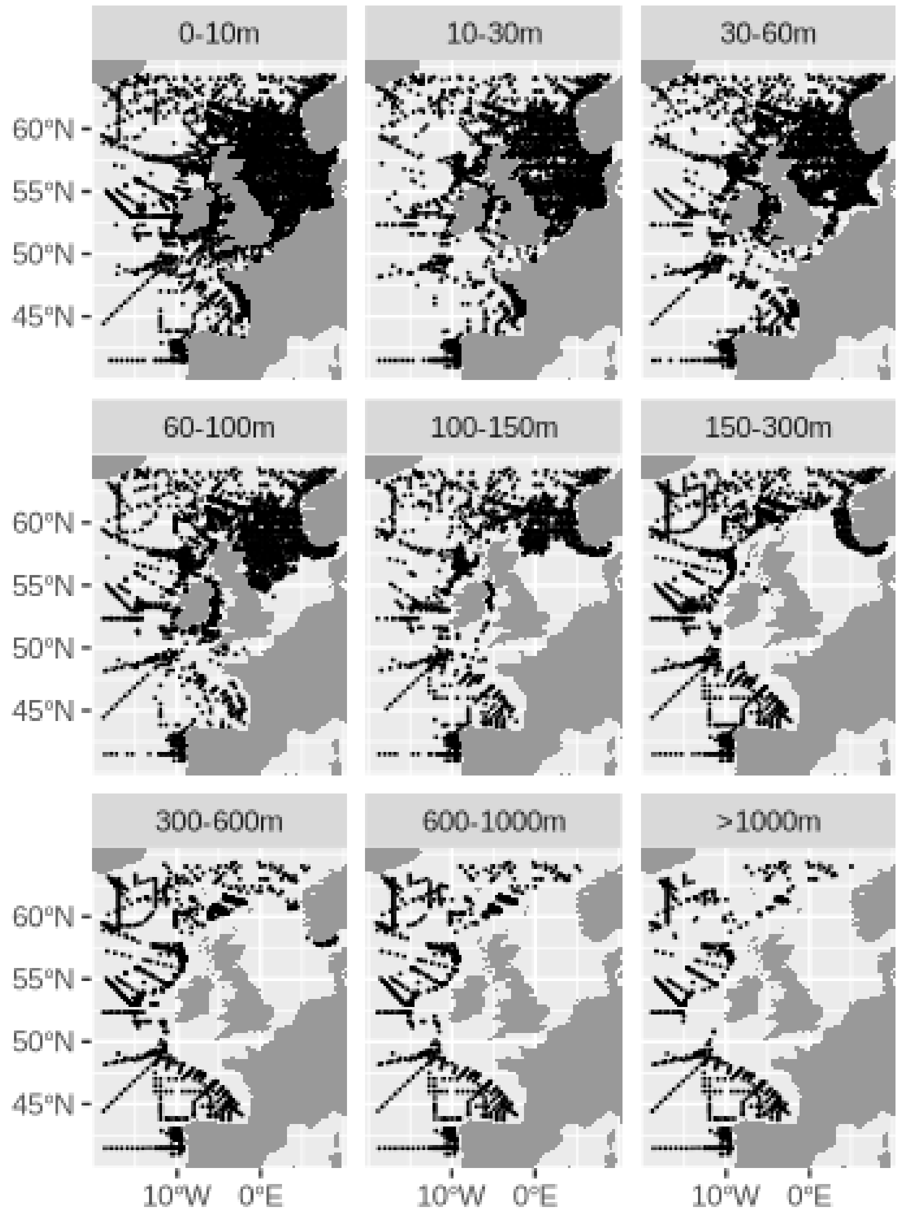


Figure 13.7: The geographic distribution of matchups between the model and observational nitrate concentration. The data has been binned into depth ranges. The depth ranges are 0-10m, 10-30m, 30-60m, 60-100m, 100-150m, 150-300m, 300-600m, 600-1000m, and >1000m. The number of observations in each depth range is shown in the tables below.

Depth	Bias	*r*	RMSD	Number of observations
0-150m	3.63	0.57	7.65	46,004
0-10m	4.51	0.60	9.93	16,147
10-30m	4.35	0.53	8.64	10,951
30-60m	3.19	0.50	4.83	7,870
60-100m	2.15	0.62	3.15	7,506
100-150m	1.54	0.61	2.52	3,530
150-300m	1.17	0.56	2.06	4,793
300-600m	0.70	0.32	2.02	4,228
600-1000m	-0.40	0.46	2.19	3,625
>1000m	-1.46	0.68	3.11	5,308

Table 13.5: Average bias (mmol Nm^{-3}), root-mean square difference (RMSD) and correlation coefficient of modelled and observed nitrate concentration for different depth ranges. The bias is calculated as model-observation. The RMSD is the square root of the mean squared difference. The correlation coefficient is the Pearson correlation coefficient between the model and observed values.

13.6 Data Sources for validation of nitrate

ICES Data Portal, Dataset on Ocean HydroChemistry, Extracted March 3, 2023. ICES, Copenhagen

CHAPTER

FOURTEEN

SEA SURFACE NITRATE VALIDATION USING GRIDDED OBSERVATIONS FROM NSBC

We used version 1.1 of the **North Sea Biogeochemical Climatology** (NSBC) to validate **sea surface nitrate**. NSBC is a monthly climatology that covers the region 47°-65°N and 15°W-15°E. The data is made up of observations over the period 1960-2014. For validation purposes we only used the level 3 data, which a gridded monthly climatology at a spatial resolution of 1/4°. The data can be download from NSBC.

Matchup procedure: The model and observations were matched up as follows. First, the model dataset was cropped by a small amount to make sure cells close to the boundary were removed. The model was then regridded to the observational grid if the observational grid was coarser using nearest neighbour. Only grid cells with model and observational data were maintained. The following model output was used to compare with the observational values: **N3_n**.

14.1 Baseline climatologies of sea surface nitrate

Climatologies of model and observational sea surface nitrate concentration are shown in the figures below. The model climatology is calculated using the years **1986-2017**.

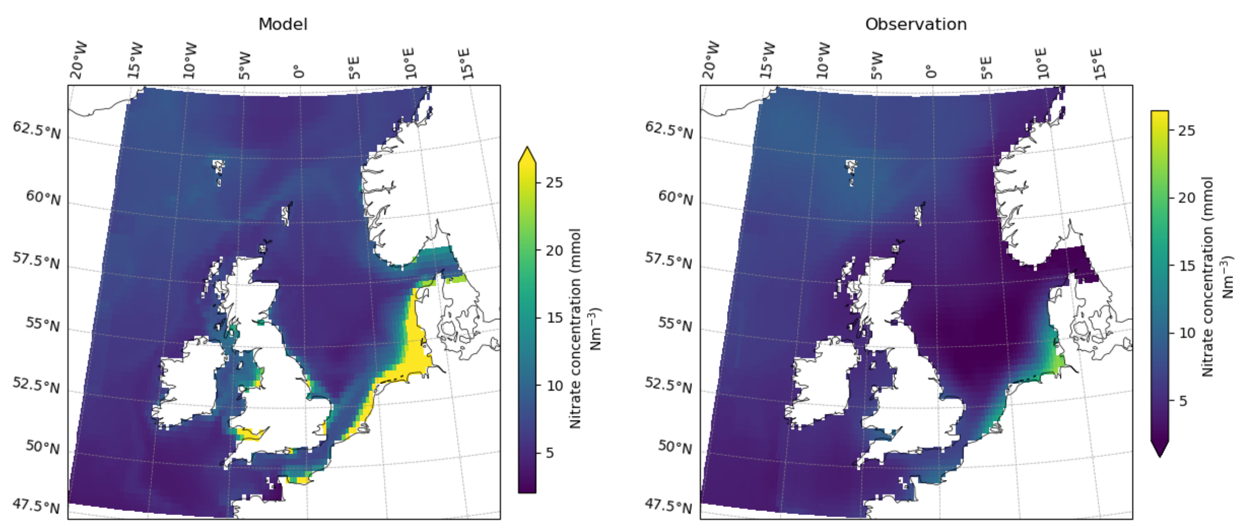


Figure 14.1: Annual average surface nitrate concentration from the model (1986-2017) and observations. Data is limited to the 2nd and 98th percentile of the combined model and observational data. Arrows indicate that values can exceed the colourbar limits.

14.2 Assessing model bias for surface nitrate concentration

Figure 14.2 shows the average bias of surface nitrate concentration simulated by the model. A positive bias indicates that the model overestimates the observation, while a negative bias indicates that the model overpredicts the observation.

The spatial average bias of surface nitrate concentration is 2.00 mmol Nm^{-3}. Overall, the model overestimates the observations in 69.3% of the model domain.

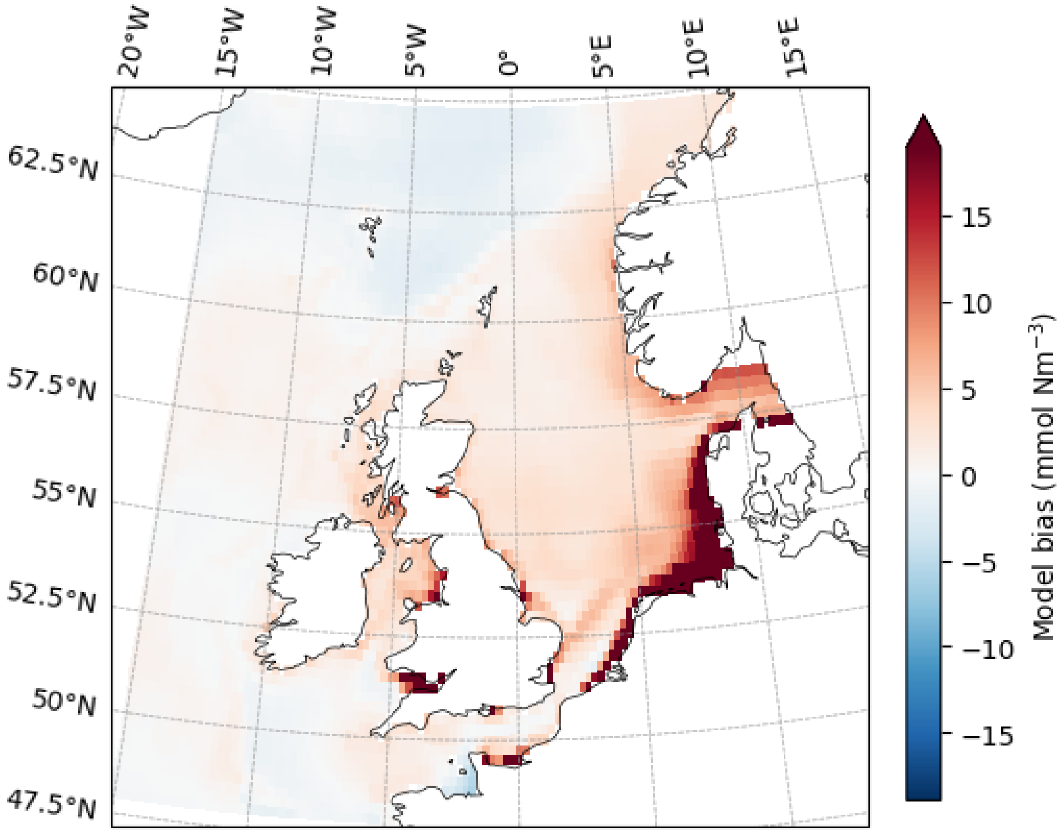


Figure 14.2: Bias of surface nitrate concentration from the model. A positive bias indicates that the model overestimates the observation. For clarity, the colourbar is limited to the 2nd and 98th percentile of the data.

14.3 Can the model reproduce seasonality of sea surface nitrate concentration?

The ability of the model to reproduce seasonality of sea surface nitrate concentration was assessed by comparing the modelled and observed seasonal cycle of nitrate concentration. First, we derive a monthly climatology for the model data. Then, we calculate the Pearson correlation coefficient between the modelled and observed nitrate concentration at each grid cell.

Note: we are only assessing the ability of the model to reproduce the ability of the model to reproduce seasonal changes, not long-term trends.

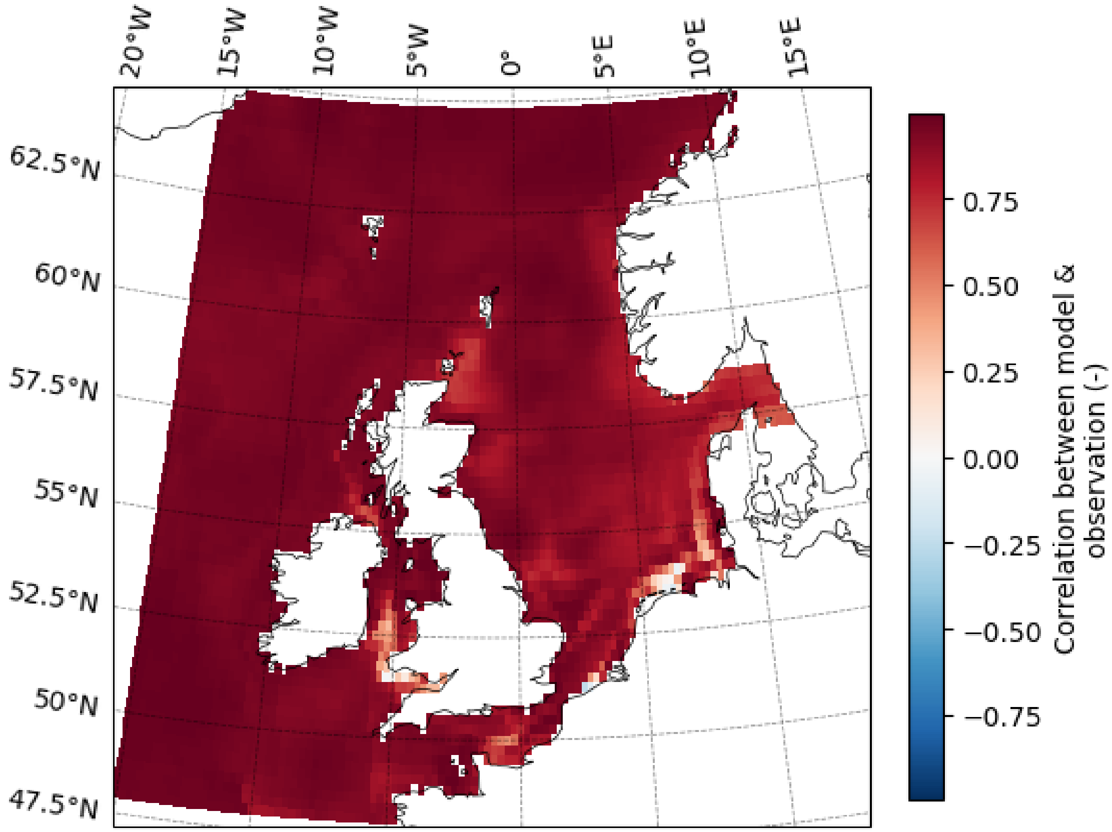
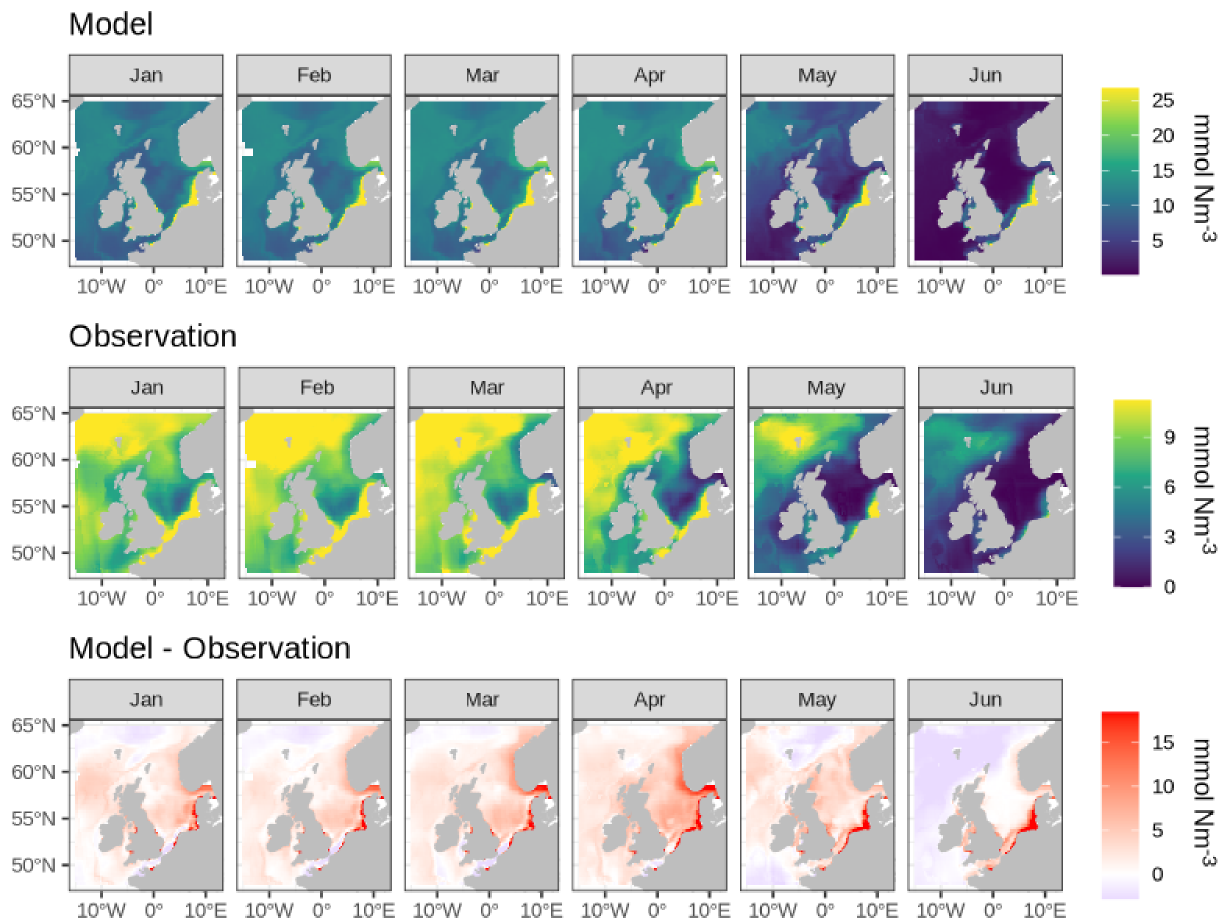


Figure 14.3: Seasonal temporal correlation between model and observations for surface nitrate concentration. This is the Pearson correlation coefficient between climatology monthly mean values in the model and observations.

The seasonal cycles of simulated and observed nitrate concentration are compared in Figure 14.4 below. This figure shows the model and observation average in each month of the year, and the differences between the two each month



Model
Jan
Feb
Mar
Apr
May
Jun
65°N
60°N
55°N
50°N
10°W
0°
10°E
25
20
15
10
5
mmol Nm-3
Observation
9
6
3
0
Model - Observation
15
10
5
0

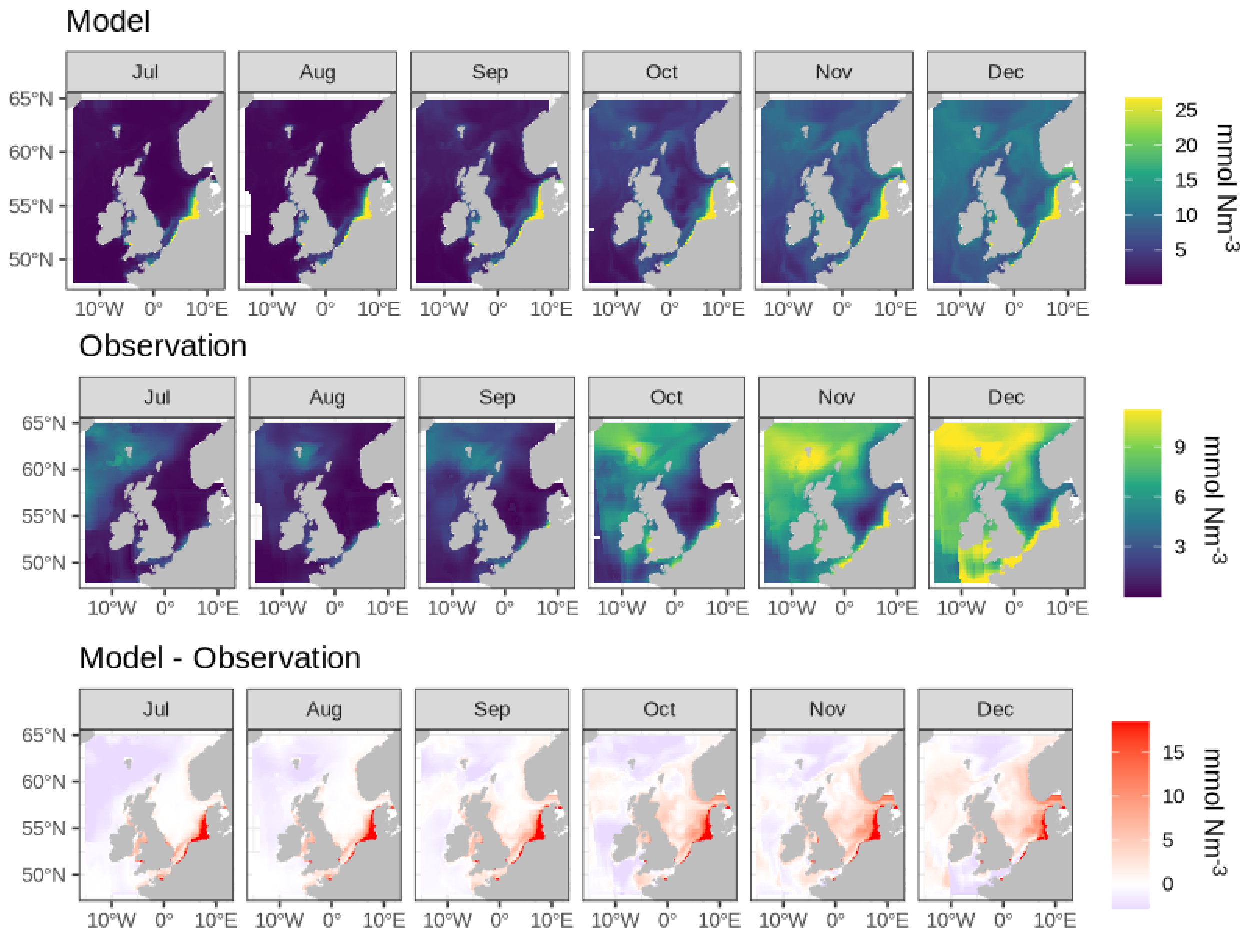


Figure 14.4: Monthly mean surface nitrate concentration for the model, observation and the difference between model and observations. For clarity, the maximum values are capped to the 98th percentiles.

14.4 Regional assessment of model performance for sea surface nitrate concentration

We assessed the regional performance of the model by comparing the model with observations in a number of regions. The regions considered are mapped below.

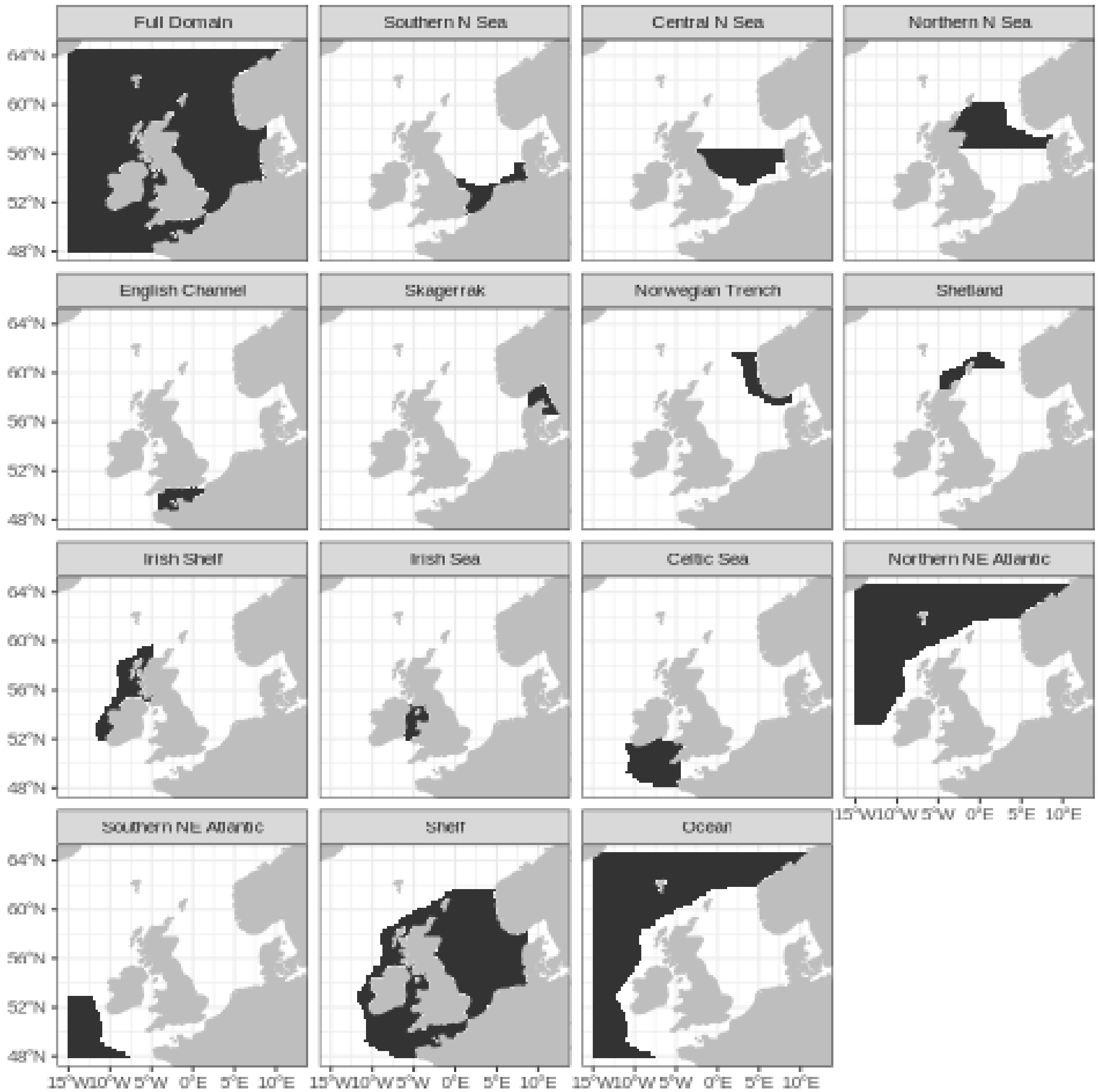


Figure 14.5: Regions used for validation of sea surface nitrate concentration.

Time series were constructed comparing the monthly mean of the spatial average sea surface nitrate concentration in each region. The spatial average was calculated using the mean of all grid cells within each region, accounting for grid cell area.

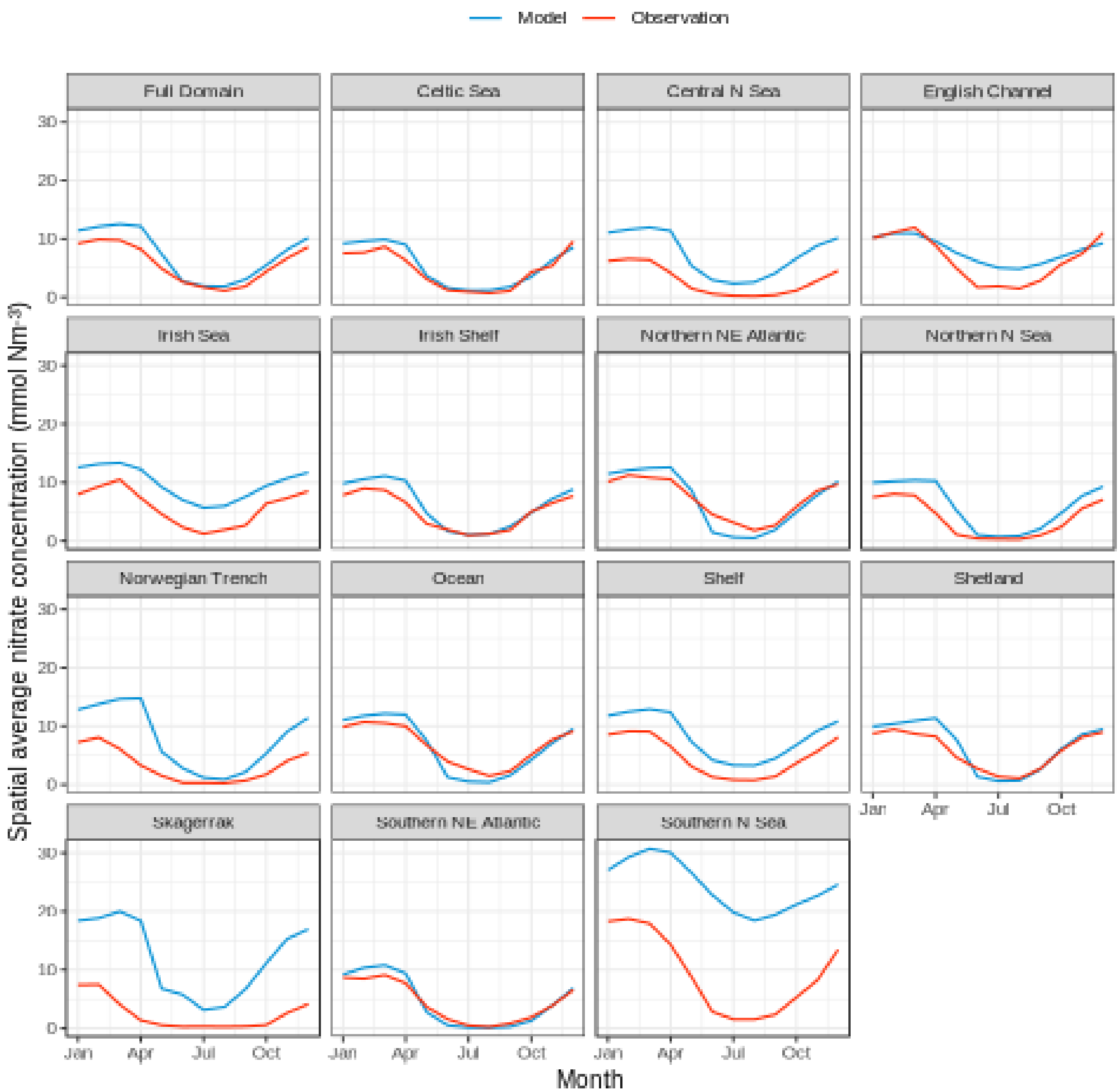


Figure 14.6: Seasonal cycle of sea surface nitrate concentration for model and observations for each region. The spatial average is taken over the region.

The table below shows the average bias of sea surface nitrate concentration in each region. The bias is calculated as the modelled value minus the observed value. A positive bias indicates that the model overestimates the observed value, while

a negative bias indicates that the model underestimates the observed value.

Region	Temporal correlation	Bias	RMSD
Full Domain	0.92	1.66	6.46
Celtic Sea	0.91	0.76	5.43
Central North Sea	0.88	4.52	6.75
Channel	0.89	1.35	5.53
Irish Sea	0.86	4.05	5.53
Irish Shelf	0.94	1.14	2.45
Northern North East Atlantic	0.95	-0.15	2.16
Northern North Sea	0.90	2.19	3.71
Norwegian Trench	0.87	4.62	5.89
Ocean	0.96	-0.08	2.05
Shelf	0.89	3.39	9.17
Shetland	0.94	0.77	1.83
Skagerrak	0.74	9.63	11.73
Southern North East Atlantic	0.97	0.21	1.32
Southern North Sea	0.77	14.95	27.59

Table 14.1: Summary of performance of the model sea surface nitrate concentration in each region. The bias (mmol Nm-3) column represents the spatial average of the annual mean modelled value minus the observed value. The temporal correlation column represents the spatial mean of the temporal correlation between the model and observations per grid cell.

14.5 Can the model reproduce spatial patterns of sea surface nitrate concentration?

The ability of the model to reproduce spatial patterns of sea surface nitrate concentration was assessed by comparing the modelled and observed nitrate concentration at each grid cell. We calculated the Pearson correlation coefficient between the modelled and observed nitrate concentration at each grid cell.

This was carried out monthly and using the annual mean in each grid cell

Time period	r
Annual mean	0.54
Jan	0.63
Feb	0.64
Mar	0.61
Apr	0.52
May	0.54
Jun	0.20
Jul	0.08
Aug	0.29
Sep	0.33
Oct	0.35
Nov	0.37
Dec	0.48

Table 14.2: Pearson correlation coefficient between modelled and observed sea surface nitrate concentration at each grid cell. The correlation was calculated monthly and using the annual mean in each grid cell.

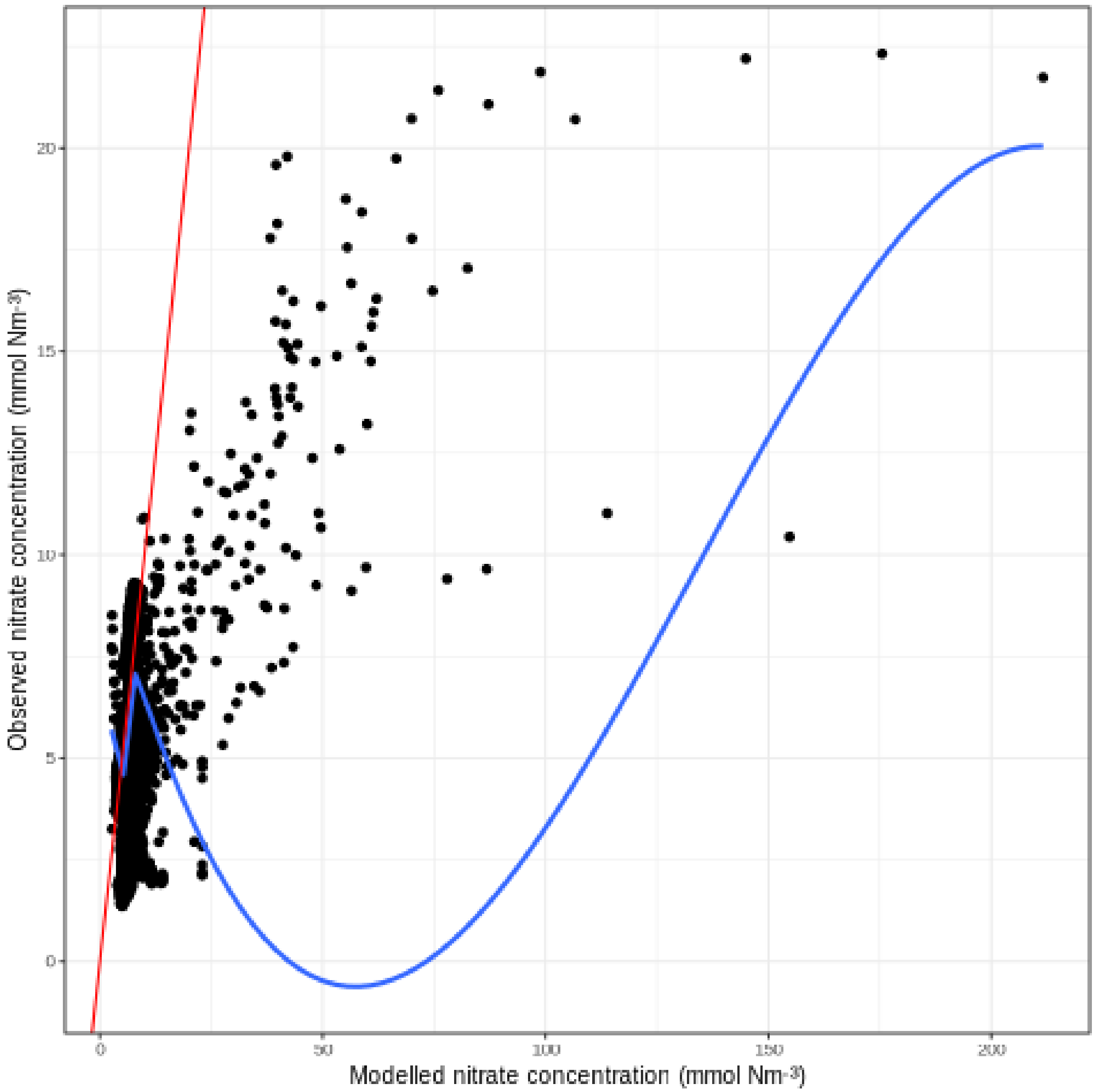


Figure 14.7: Scatter plot of modelled and observed annual average sea surface nitrate concentration in grid cells. The red line is the 1:1 line. The blue line is the linear regression of the modelled and observed values. The slope of the blue line is the slope of the regression line.

14.6 Data Sources for validation of nitrate concentration

Hinrichs, Iris; Gouretski, Viktor; Paetsch, Johannes; Emeis, Kay; Stammer, Detlef (2017). North Sea Biogeochemical Climatology (Version 1.1).

URL: https://www.cen.uni-hamburg.de/en/icdc/data/ocean/nsbc.html

CHAPTER

FIFTEEN

VALIDATION OF OXYGEN USING POINT OBSERVATIONS

15.1 Performance of model sea surface oxygen

Values from the **top 5m** of the water column were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 27,851 values extracted from the observational database. The map below shows the locations of the matched up data for oxygen concentration.

The following model output was used to compare with observational values: **O2_o**.

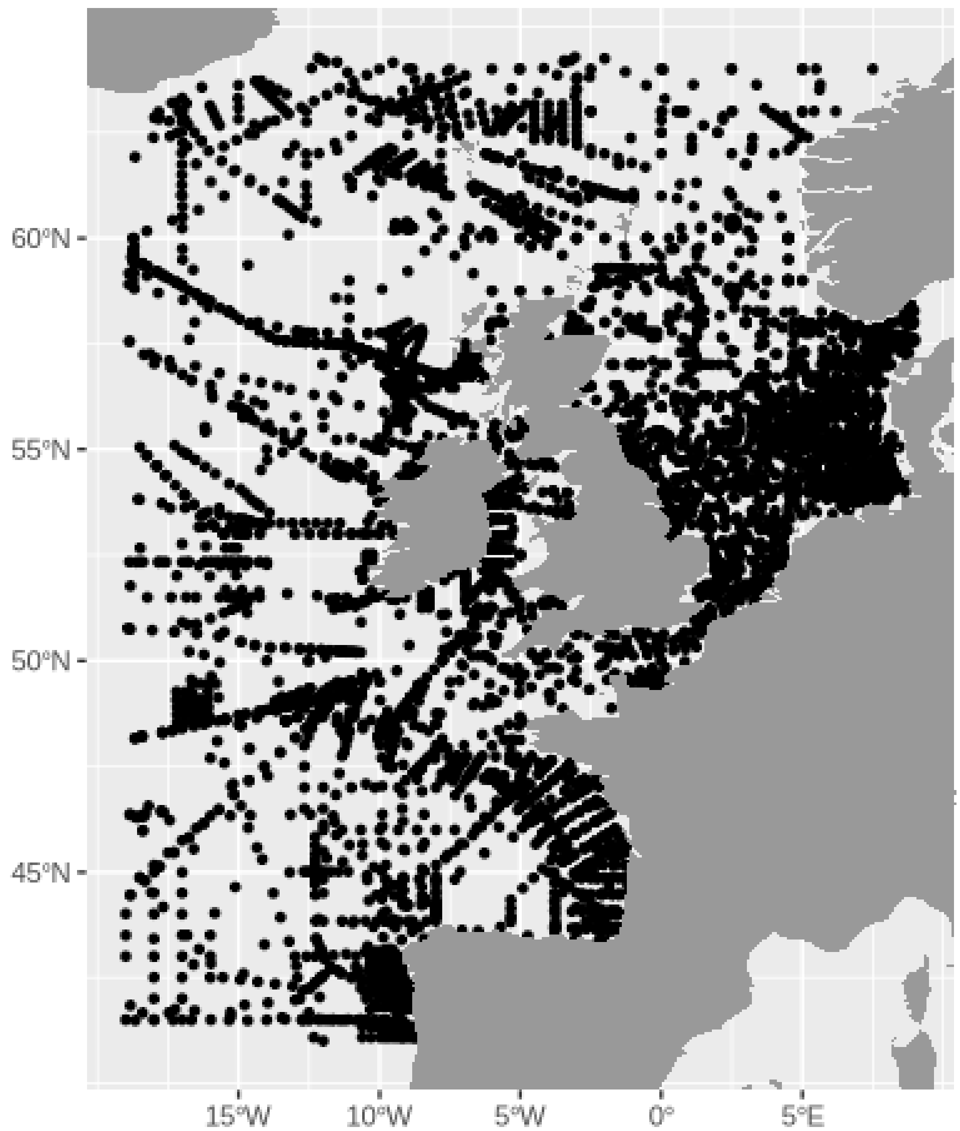


Figure 15.1: Locations of matchups between simulated and observed oxygen concentration in the top 5m of the water column.

The number of observations in each month ranged from 892 in December to 5,025 in August. Figure 15.2 below shows the distribution of observations in each month.

Figure 15.2 below shows the bias between the model and observational data for oxygen concentration. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

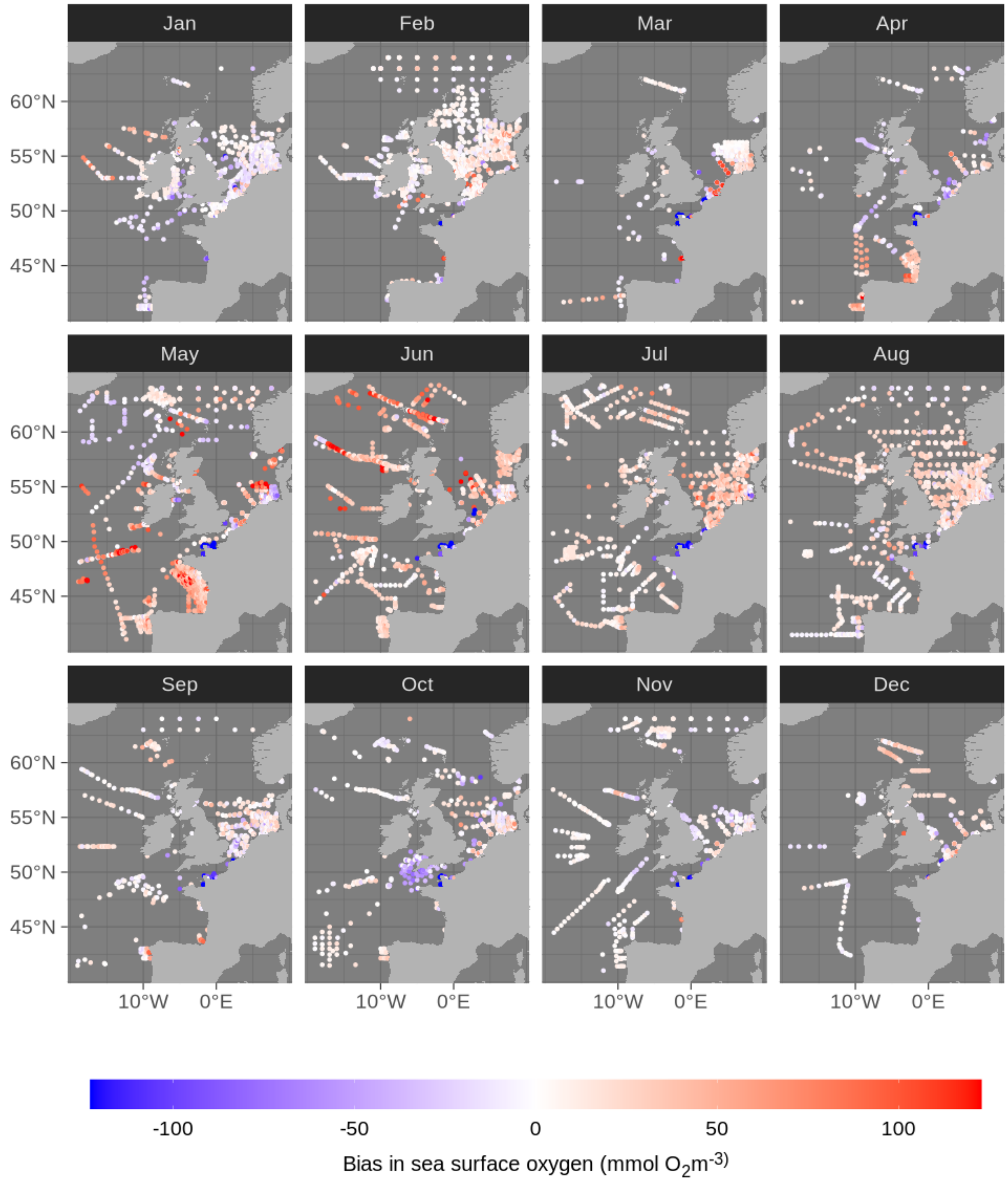


Figure 15.2: Bias in sea surface oxygen concentration. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 15.3 and 15.4 show the distribution of sea surface oxygen concentration observations in the model and observational datasets. This is shown for each month of the year (Figure 15.3) and for the entire year (Figure 15.4).

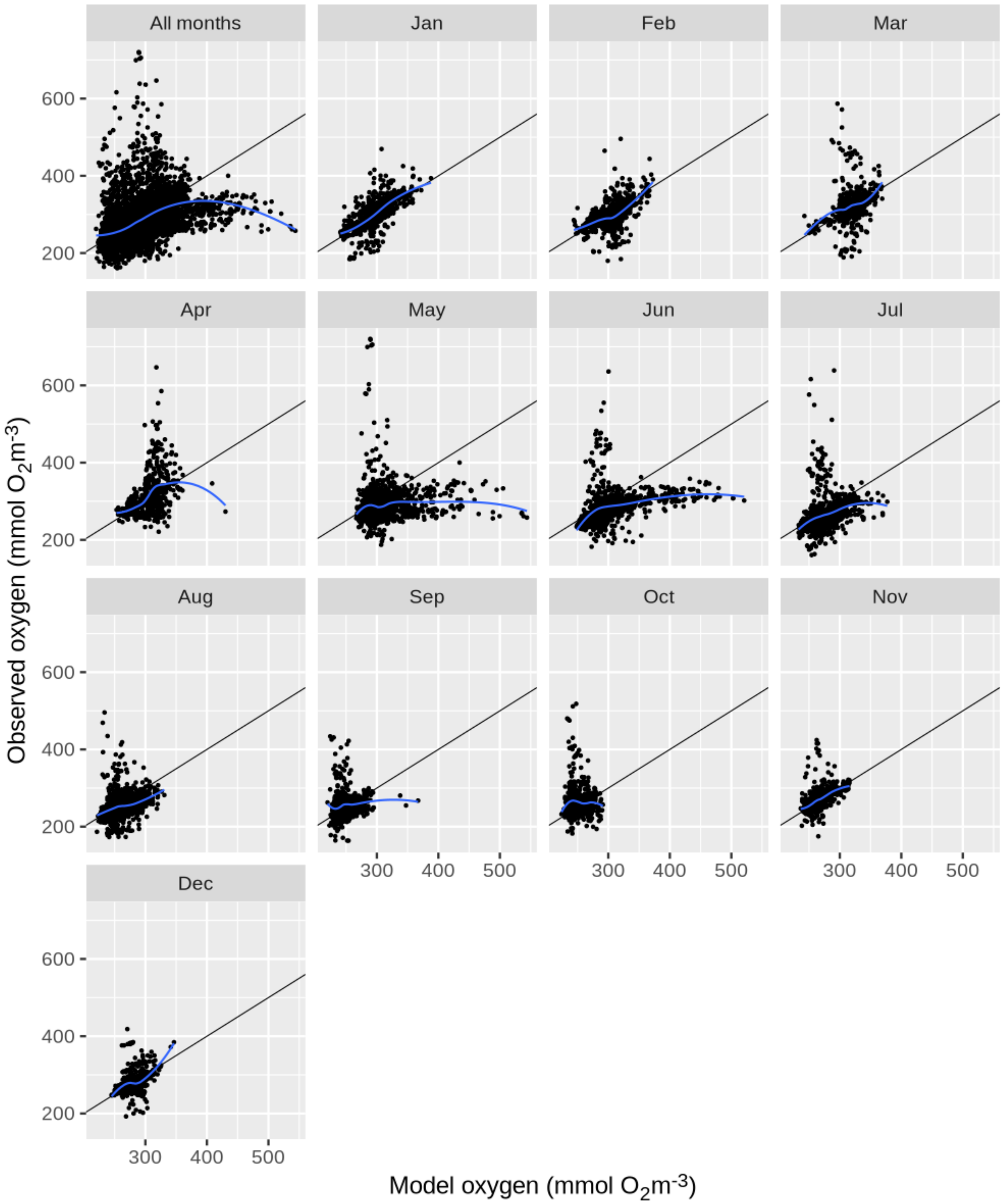


Figure 15.3: Simulated versus observed oxygen concentration in the top 5m of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

15.2 Summary statistics for sea surface oxygen

The overall ability of the model to predict the observed oxygen concentration was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	9.62	37.59	0.50	27,851
Jan	-8.80	26.46	0.68	2,068
Feb	11.97	26.48	0.42	3,608
Mar	0.10	39.37	0.40	1,084
Apr	-7.77	45.23	0.39	1,399
May	29.00	56.68	0.16	3,561
Jun	20.87	48.10	0.41	2,558
Jul	15.42	37.84	0.36	2,780
Aug	11.77	28.19	0.28	5,025
Sep	0.46	34.68	0.18	1,558
Oct	-2.04	38.95	-0.12	1,583
Nov	-2.20	19.27	0.63	1,735
Dec	1.53	25.52	0.41	892

Table 15.1: Average bias (mmol $O_2 m^{-3}$) and root-mean square deviation (mmol $O_2 m^{-3}$) for the model's sea surface oxygen concentration for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed oxygen concentration was performed. The modelled oxygen concentration was used as the independent variable and the observed oxygen concentration was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.61	105.61	0.22	0.00
Jan	1.10	-25.61	0.49	0.00
Feb	0.79	55.50	0.29	0.00
Mar	0.74	86.32	0.10	0.00
Apr	0.98	16.77	0.15	0.00
May	0.11	256.00	0.01	0.00
Jun	0.36	174.11	0.13	0.00
Jul	0.58	104.94	0.10	0.00
Aug	0.46	132.97	0.09	0.00
Sep	0.33	172.42	0.03	0.00
Oct	-0.06	277.60	0.00	0.59
Nov	0.88	34.44	0.31	0.00
Dec	0.75	69.41	0.16	0.00

Table 15.2: Linear regression analysis of modelled and observed oxygen. The modelled oxygen concentration was used as the independent variable and the observed oxygen concentration was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

15.3 Performance of model near-bottom oxygen

Near-bottom values of oxygen were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The near-bottom was defined as observations **within 2m of the seabed**. This was interpolated to the observational grid using the $GEBCO_2$ bathymetry dataset. Model values were interpolated to the observational dataset's longitudes and latitudes using 3D interpolation. **Note:** this analysis has been restricted to observations on the shelf region. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 6,045 values extracted from the observational database. The map below shows the locations of the matched up data for oxygen concentration.

The following model output was used to compare with observational values: **O2_o**.

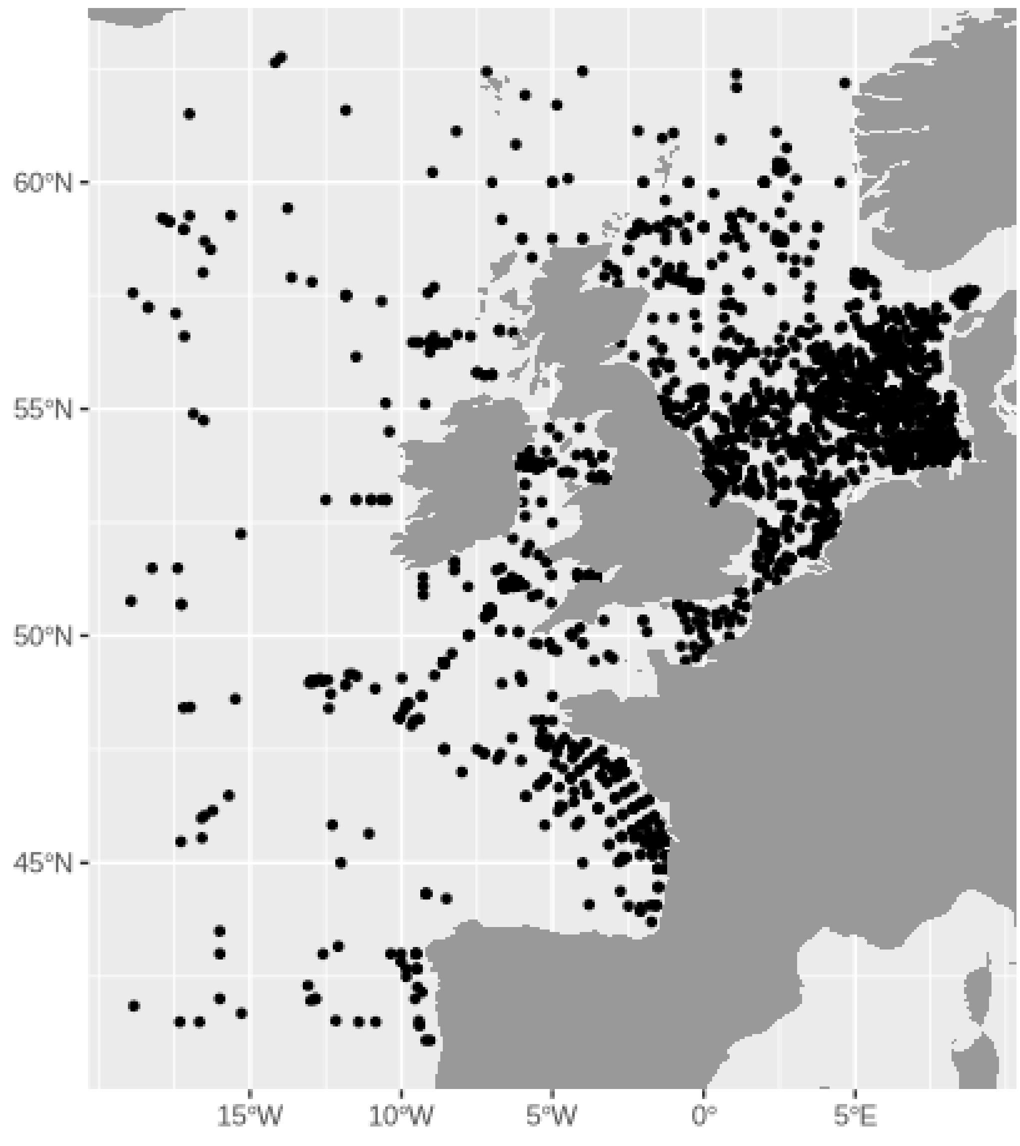


Figure 15.4: Locations of matchups between simulated and observed oxygen concentration near the bottom of the water

column.

The number of observations in each month ranged from 100 in December to 1,474 in February. Figure 15.5 below shows the distribution of observations in each month.

Figure 15.5 below shows the bias between the model and observational data for oxygen concentration. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

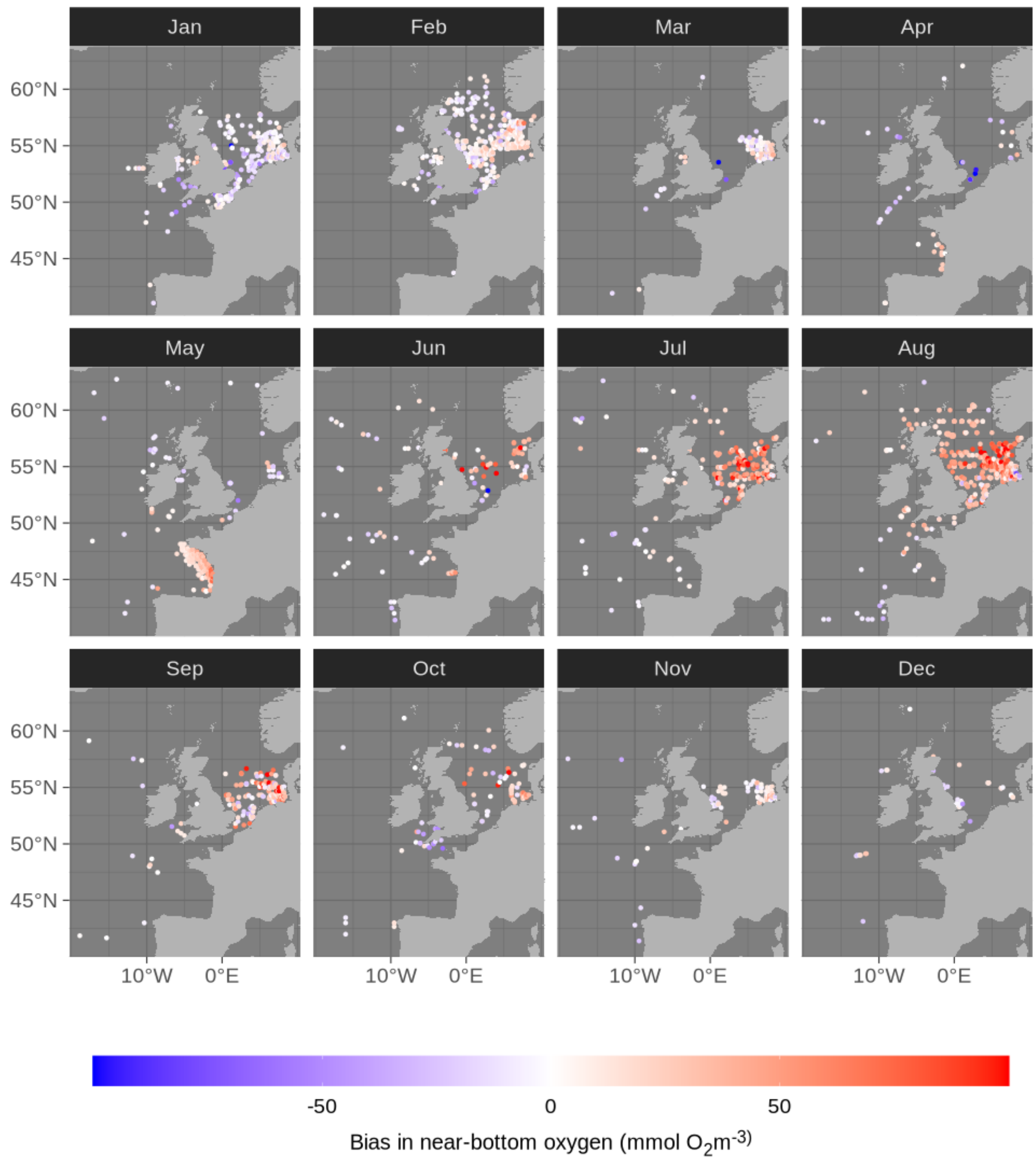


Figure 15.5: Bias in near-bottom oxygen concentration. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 15.6 and 15.7 show the distribution of near-bottom oxygen concentration observations in the model and observational datasets. This is shown for each month of the year (Figure 15.6) and for the entire year (Figure 15.7).

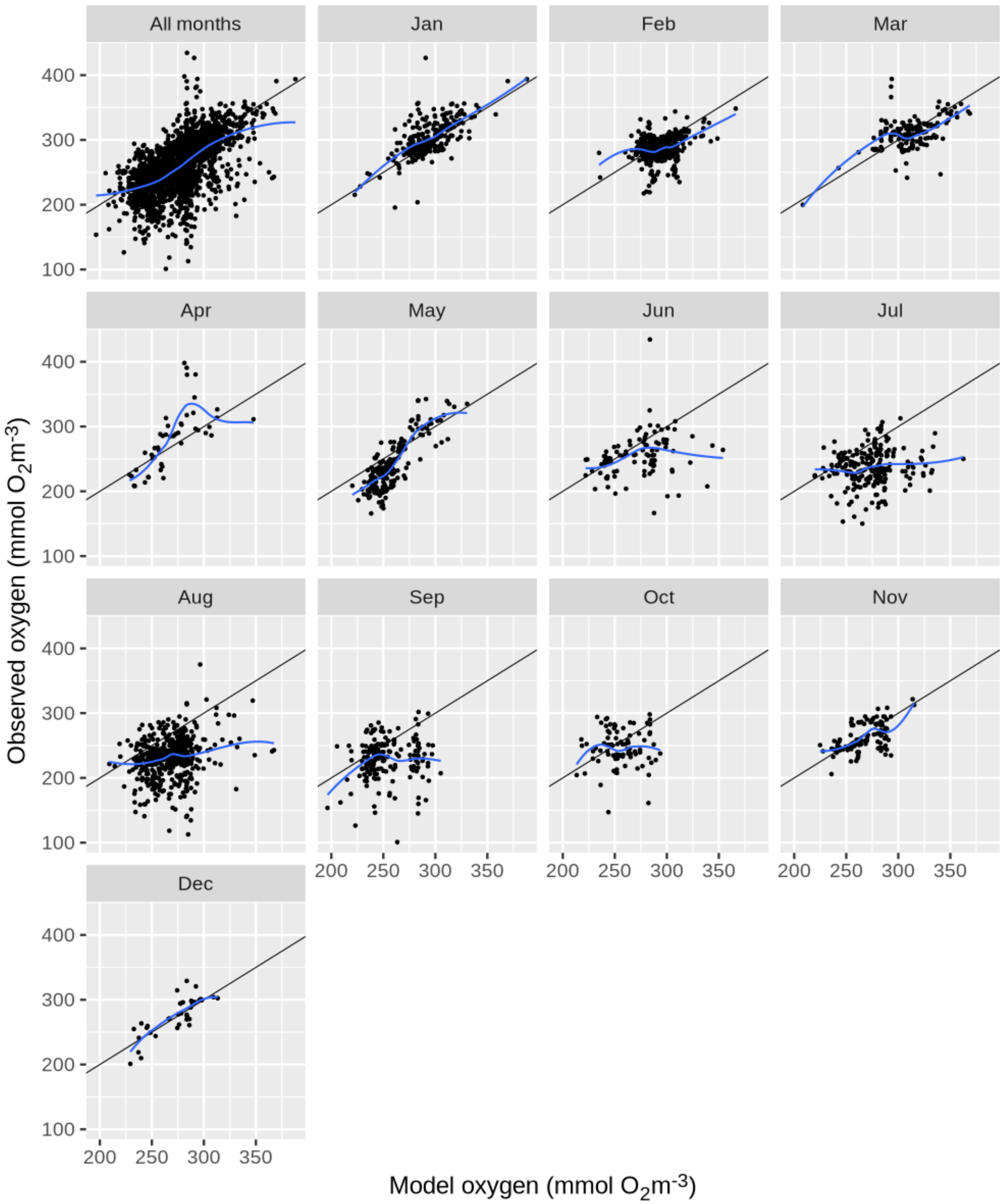


Figure 15.6: Simulated versus observed oxygen concentration near the bottom of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

15.4 Summary statistics for near-bottom oxygen

The overall ability of the model to predict the observed oxygen concentration was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	14.72	35.54	0.62	6,045
Jan	-13.02	27.26	0.56	542
Feb	6.43	19.54	0.25	1,474
Mar	-2.89	24.37	0.54	339
Apr	-15.43	56.56	0.52	136
May	20.03	31.24	0.77	614
Jun	16.96	42.89	0.25	227
Jul	35.77	47.67	0.04	554
Aug	33.91	45.76	0.17	1,349
Sep	25.54	46.25	0.24	276
Oct	5.60	32.27	0.14	185
Nov	1.38	21.81	0.52	249
Dec	-4.83	18.23	0.75	100

Table 15.3: Average bias (mmol $O_2 m^{-3}$) and root-mean square deviation (mmol $O_2 m^{-3}$) for the model's near-bottom oxygen concentration for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed oxygen concentration was performed. The modelled oxygen concentration was used as the independent variable and the observed oxygen concentration was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.94	0.77	0.37	0.00
Jan	0.91	34.17	0.46	0.00
Feb	0.37	177.80	0.09	0.00
Mar	0.54	141.71	0.26	0.00
Apr	1.29	-66.71	0.46	0.00
May	1.60	-172.05	0.68	0.00
Jun	0.30	173.12	0.06	0.01
Jul	0.15	196.49	0.02	0.07
Aug	0.32	145.83	0.05	0.00
Sep	0.14	194.61	0.01	0.20
Oct	0.07	227.24	0.00	0.66
Nov	0.64	90.07	0.29	0.00
Dec	1.05	-11.76	0.66	0.00

Table 15.4: Linear regression analysis of modelled and observed oxygen. The modelled oxygen concentration was used as the independent variable and the observed oxygen concentration was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

15.5 Depth-resolved comparisons of modelled and observed oxygen

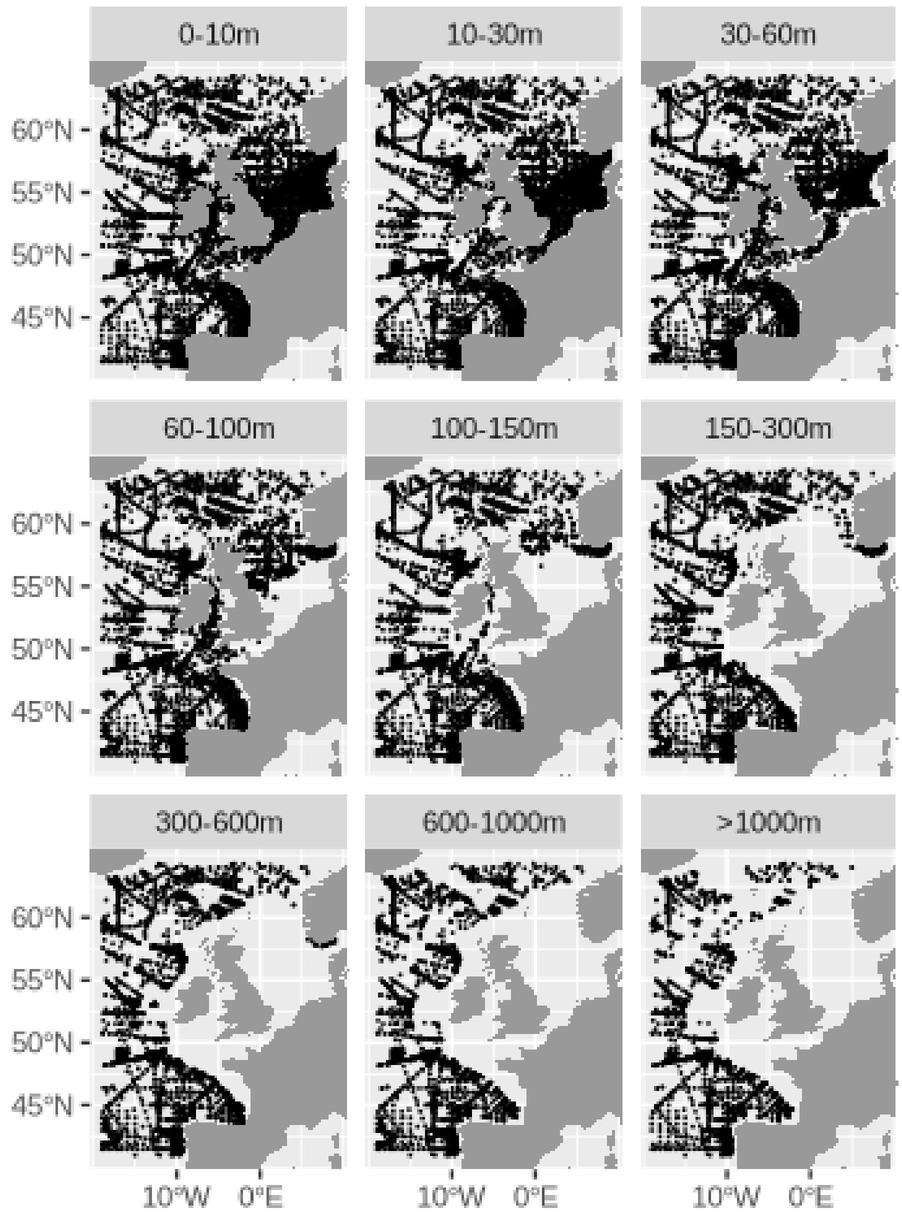


Figure 15.7: The geographic distribution of matchups between the model and observational oxygen concentration. The data has been binned into depth ranges. The depth ranges are 0-10m, 10-30m, 30-60m, 60-100m, 100-150m, 150-300m, 300-600m, 600-1000m, and >1000m. The number of observations in each depth range is shown in the tables below.

Depth	Bias	*r*	RMSD	Number of observations
0-150m	10.41	0.55	25.88	488,458
0-10m	14.77	0.57	31.49	49,764
10-30m	18.30	0.50	33.63	106,371
30-60m	14.71	0.38	30.32	116,269
60-100m	4.73	0.60	17.31	113,034
100-150m	1.51	0.71	12.79	103,020
150-300m	-0.05	0.79	11.29	227,657
300-600m	-0.25	0.82	13.29	372,786
600-1000m	16.59	0.77	25.04	401,111
>1000m	-5.64	0.65	17.24	1,638,137

Table 15.5: Average bias (mmol $O_2 m^{-3}$), root-mean square difference (RMSD) and correlation coefficient of modelled and observed oxygen concentration for different depth ranges. The bias is calculated as model-observation. The RMSD is the square root of the mean squared difference. The correlation coefficient is the Pearson correlation coefficient between the model and observed values.

15.6 Data Sources for validation of oxygen

ICES Data Portal, Dataset on Ocean HydroChemistry, Extracted March 3, 2023. ICES, Copenhagen

CHAPTER

SIXTEEN

SEA SURFACE OXYGEN VALIDATION USING GRIDDED OBSERVATIONS FROM NSBC

We used version 1.1 of the **North Sea Biogeochemical Climatology** (NSBC) to validate **sea surface oxygen**. NSBC is a monthly climatology that covers the region 47°-65°N and 15°W-15°E. The data is made up of observations over the period 1960-2014. For validation purposes we only used the level 3 data, which a gridded monthly climatology at a spatial resolution of 1/4°. The data can be download from NSBC.

Matchup procedure: The model and observations were matched up as follows. First, the model dataset was cropped by a small amount to make sure cells close to the boundary were removed. The model was then regridded to the observational grid if the observational grid was coarser using nearest neighbour. Only grid cells with model and observational data were maintained. The following model output was used to compare with the observational values: **O2_o**.

16.1 Baseline climatologies of sea surface oxygen

Climatologies of model and observational sea surface oxygen concentration are shown in the figures below. The model climatology is calculated using the years **1986-2017**.

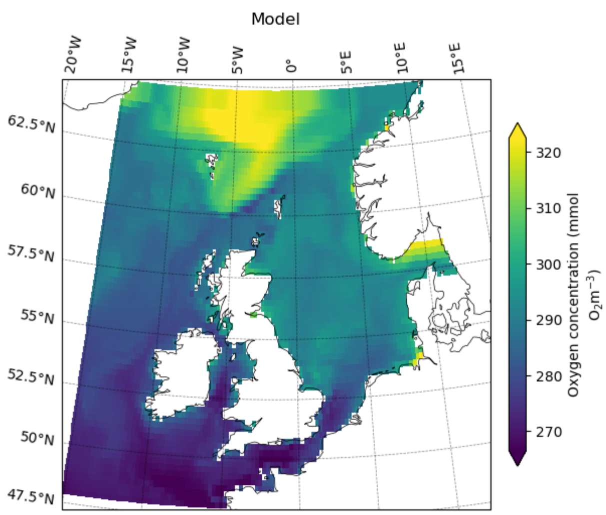


Figure 16.1: Annual average surface oxygen concentration from the model (1986-2017) and observations. Data is limited to the 2^{nd} and 98^{th} percentile of the combined model and observational data. Arrows indicate that values can exceed the colourbar limits.

16.2 Assessing model bias for surface oxygen concentration

Figure 16.2 shows the average bias of surface oxygen concentration simulated by the model. A positive bias indicates that the model overestimates the observation, while a negative bias indicates that the model overpredicts the observation.

The spatial average bias of surface oxygen concentration is 4.64 mmol O_2m^{-3}. Overall, the model overestimates the observations in 79.3% of the model domain.

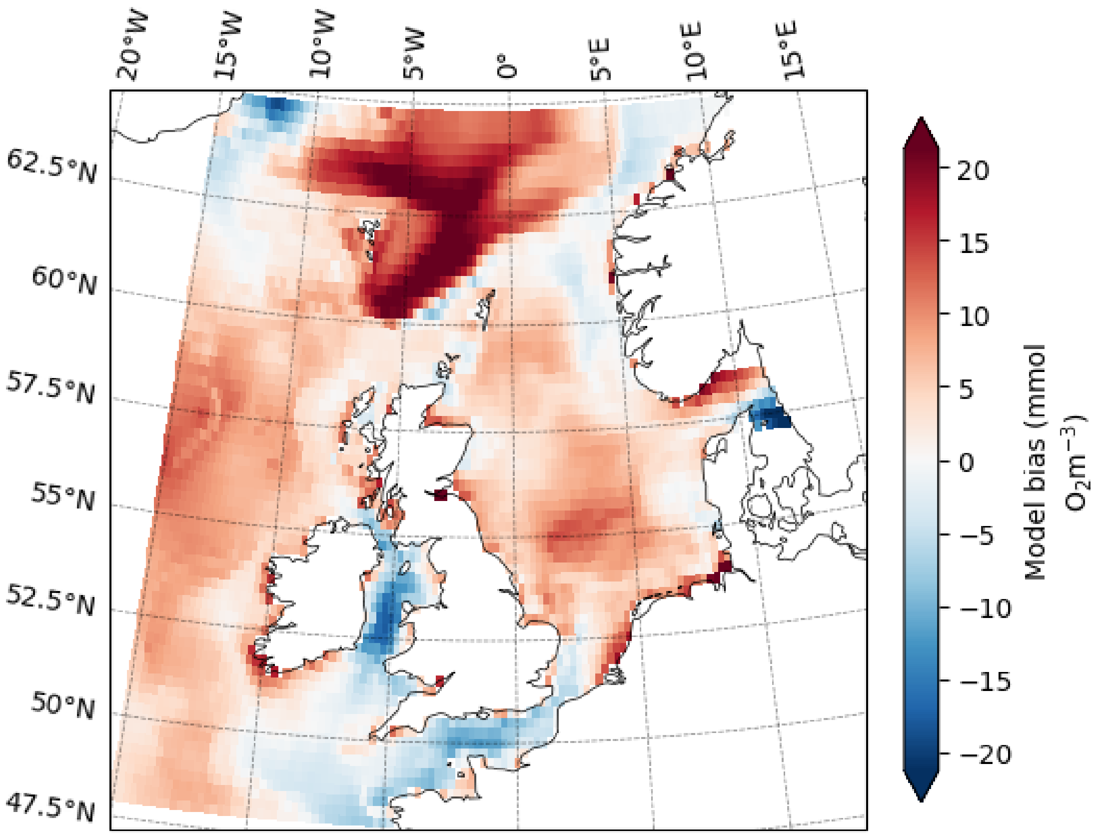


Figure 16.2: Bias of surface oxygen concentration from the model. A positive bias indicates that the model overestimates the observation. For clarity, the colourbar is limited to the 2nd and 98th percentile of the data.

16.3 Can the model reproduce seasonality of sea surface oxygen concentration?

The ability of the model to reproduce seasonality of sea surface oxygen concentration was assessed by comparing the modelled and observed seasonal cycle of oxygen concentration. First, we derive a monthly climatology for the model data. Then, we calculate the Pearson correlation coefficient between the modelled and observed oxygen concentration at each grid cell.

Note: we are only assessing the ability of the model to reproduce the ability of the model to reproduce seasonal changes, not long-term trends.

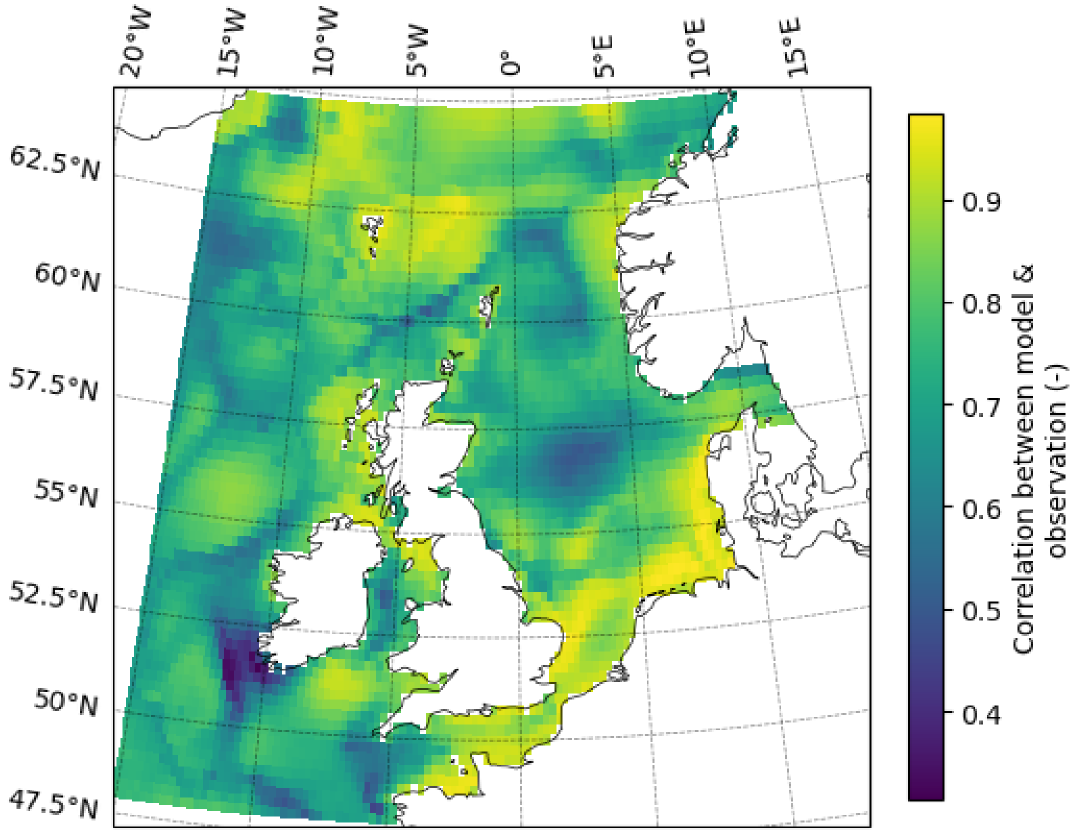
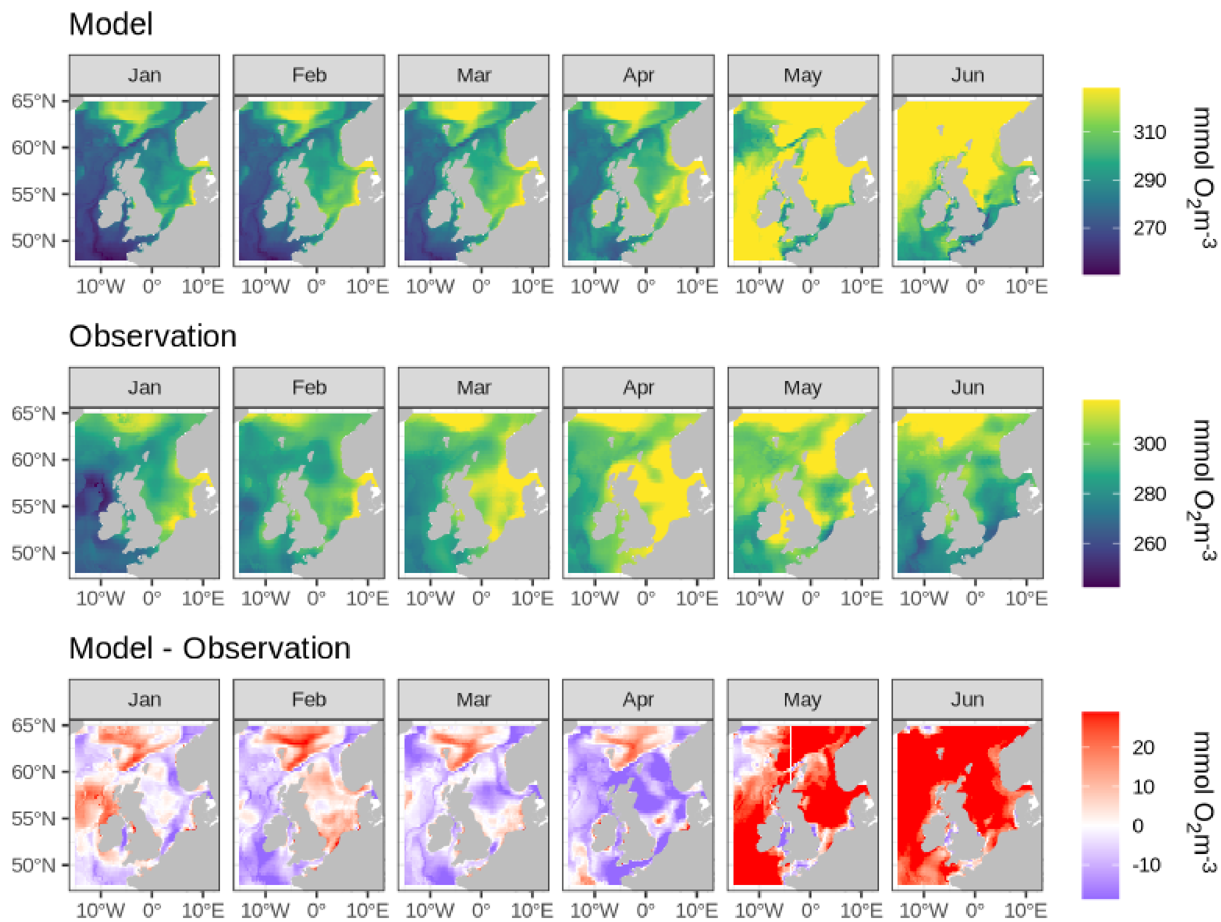


Figure 16.3: Seasonal temporal correlation between model and observations for surface oxygen concentration. This is the Pearson correlation coefficient between climatology monthly mean values in the model and observations.

The seasonal cycles of simulated and observed oxygen concentration are compared in Figure 16.4 below. This figure shows the model and observation average in each month of the year, and the differences between the two each month


Model
Jan
Feb
Mar
Apr
May
Jun
65°N
60°N
55°N
50°N
10°W
0°
10°E
310
290
270
mmol O2m-3
Observation
300
280
260
Model - Observation
20
10
0
-10

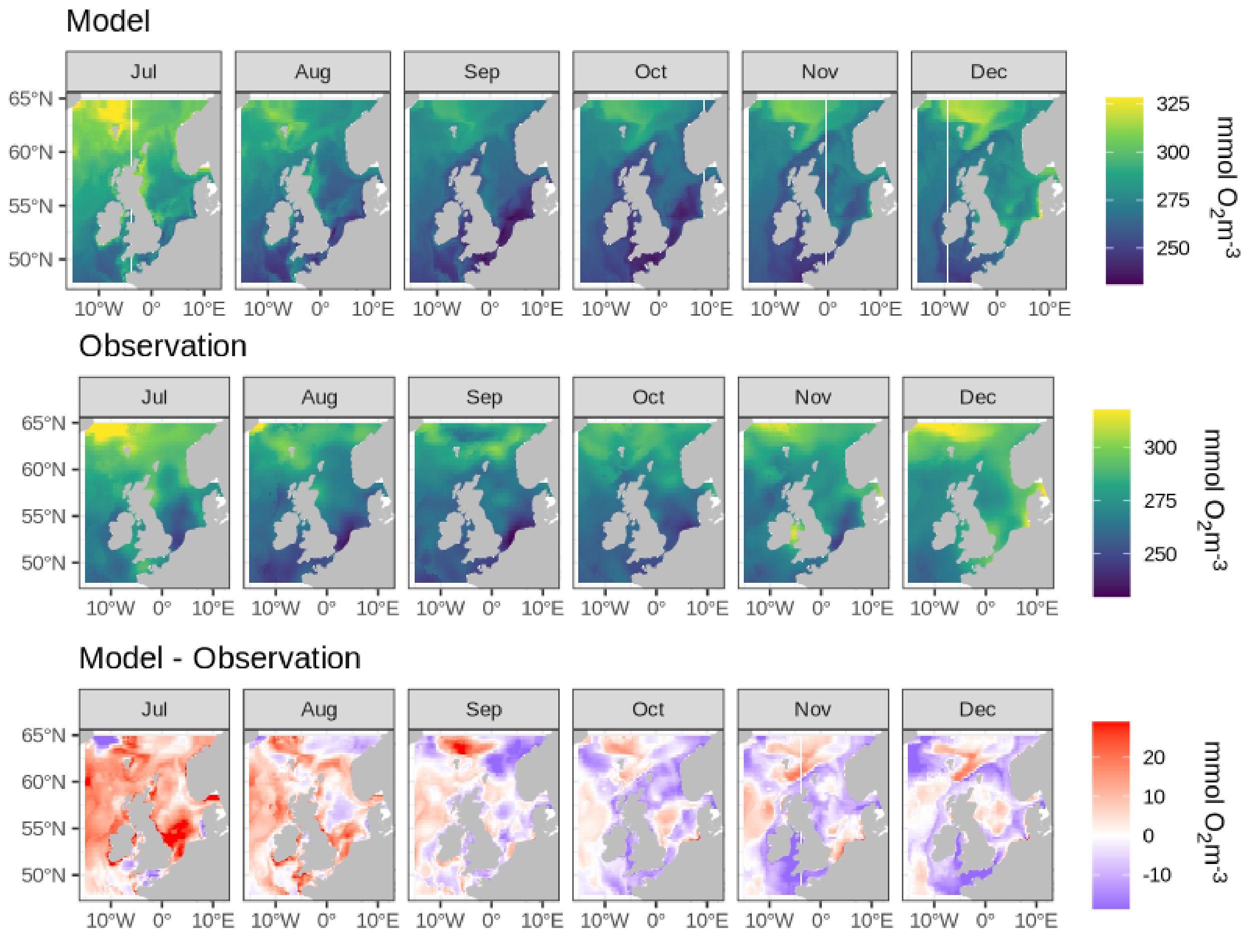


Figure 16.4: Monthly mean surface oxygen concentration for the model, observation and the difference between model and observations. For clarity, the maximum values are capped to the 98th percentiles.

16.4 Regional assessment of model performance for sea surface oxygen concentration

We assessed the regional performance of the model by comparing the model with observations in a number of regions. The regions considered are mapped below.

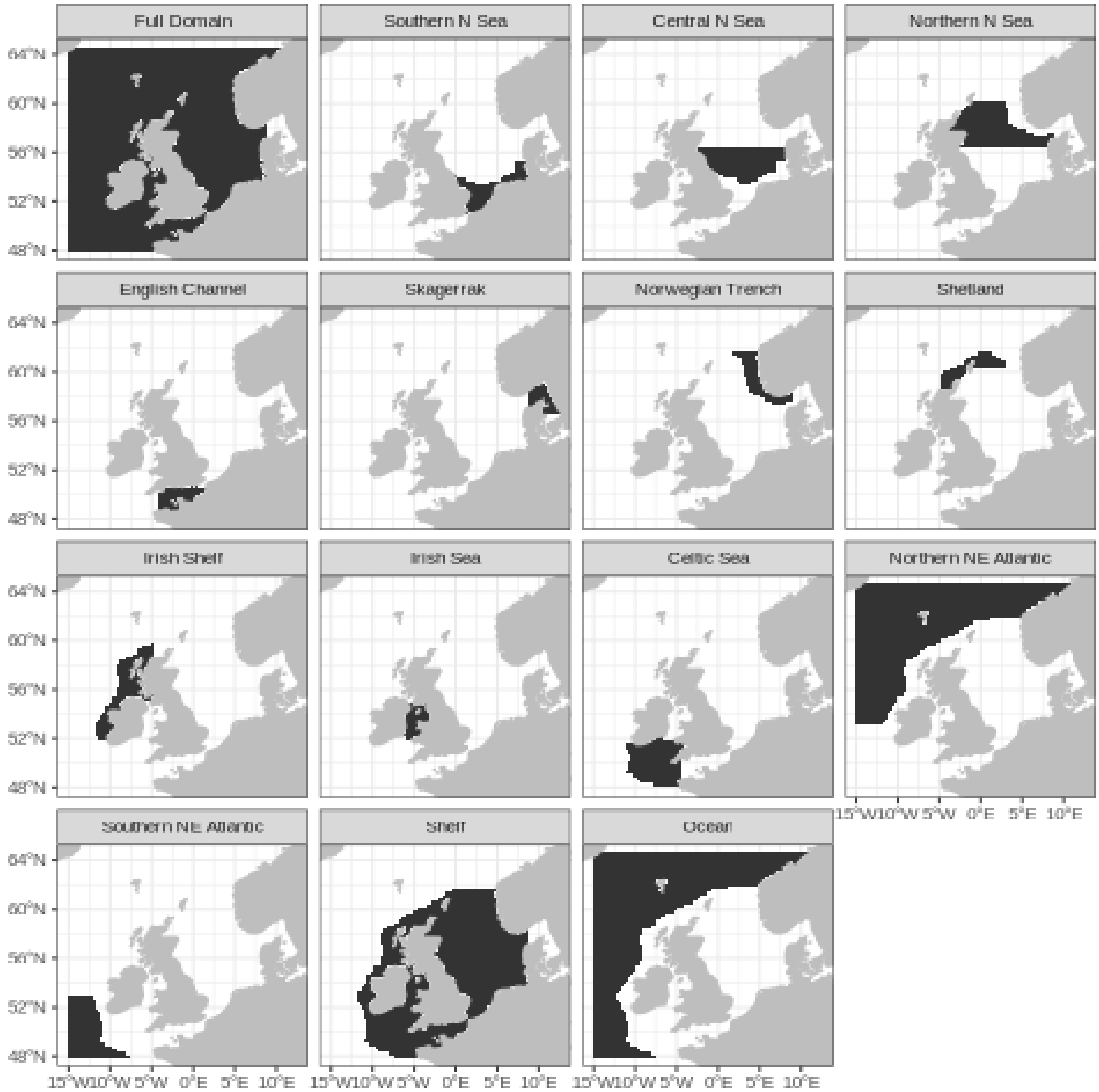


Figure 16.5: Regions used for validation of sea surface oxygen concentration.

Time series were constructed comparing the monthly mean of the spatial average sea surface oxygen concentration in each region. The spatial average was calculated using the mean of all grid cells within each region, accounting for grid cell area.

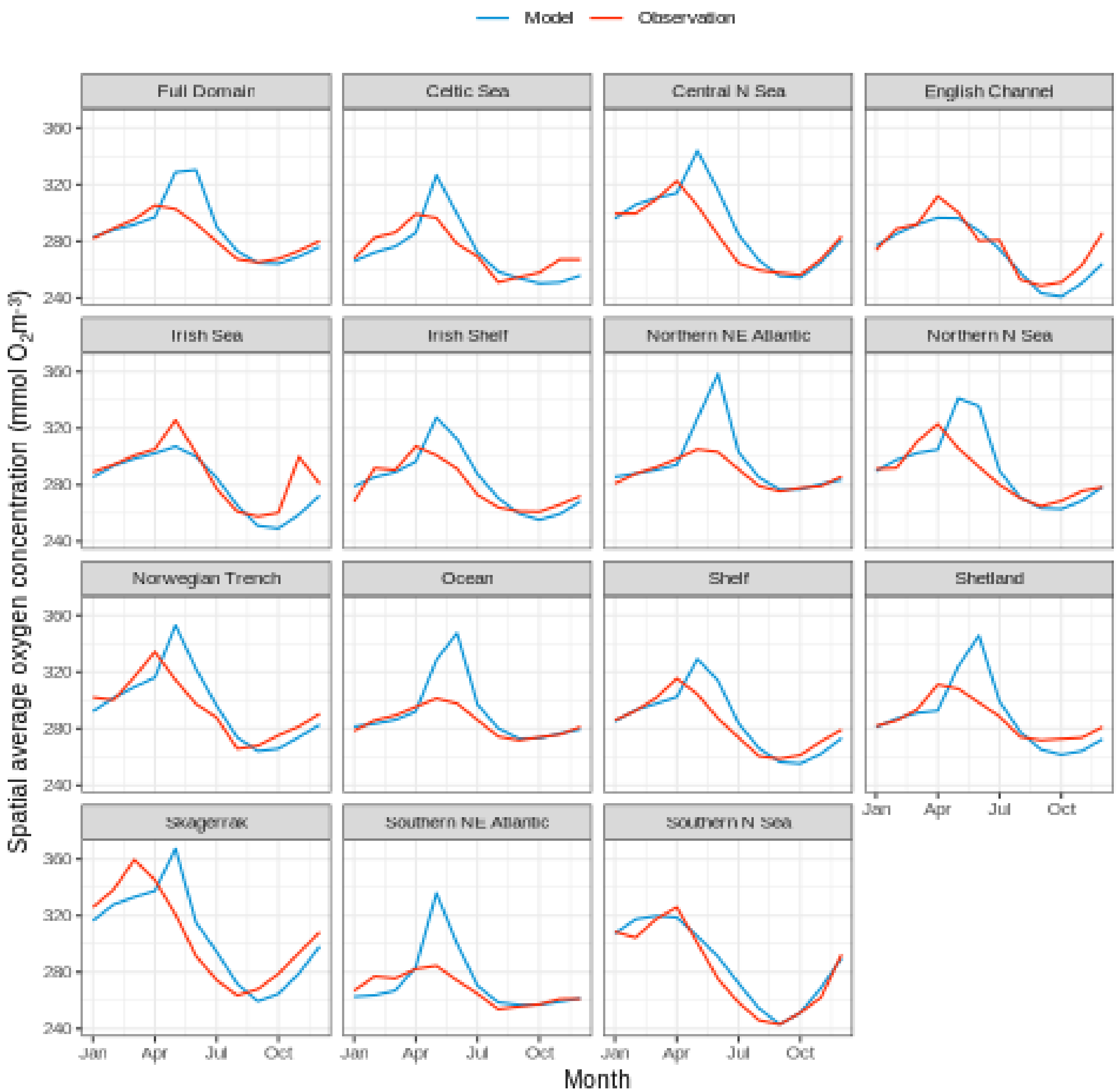


Figure 16.6: Seasonal cycle of sea surface oxygen concentration for model and observations for each region. The spatial average is taken over the region.

The table below shows the average bias of sea surface oxygen concentration in each region. The bias is calculated as the modelled value minus the observed value. A positive bias indicates that the model overestimates the observed value, while

a negative bias indicates that the model underestimates the observed value.

Region	Temporal correlation	Bias	RMSD
Full Domain	0.76	4.70	18.97
Celtic Sea	0.72	-0.77	16.39
Central North Sea	0.80	7.00	18.63
Channel	0.88	-5.36	12.17
Irish Sea	0.75	-7.12	17.99
Irish Shelf	0.77	3.61	14.77
Northern North East Atlantic	0.77	7.76	20.93
Northern North Sea	0.71	4.44	19.32
Norwegian Trench	0.77	1.48	17.36
Ocean	0.76	7.28	20.52
Shelf	0.77	2.23	17.15
Shetland	0.74	1.80	17.94
Skagerrak	0.80	-0.26	24.77
Southern North East Atlantic	0.71	5.21	18.04
Southern North Sea	0.92	4.47	14.13

Table 16.1: Summary of performance of the model sea surface oxygen concentration in each region. The bias (mmol O2m-3) column represents the spatial average of the annual mean modelled value minus the observed value. The temporal correlation column represents the spatial mean of the temporal correlation between the model and observations per grid cell.

16.5 Can the model reproduce spatial patterns of sea surface oxygen concentration?

The ability of the model to reproduce spatial patterns of sea surface oxygen concentration was assessed by comparing the modelled and observed oxygen concentration at each grid cell. We calculated the Pearson correlation coefficient between the modelled and observed oxygen concentration at each grid cell.

This was carried out monthly and using the annual mean in each grid cell

Time period	*r*
Annual mean	0.87
Jan	0.84
Feb	0.83
Mar	0.87
Apr	0.74
May	0.17
Jun	0.79
Jul	0.80
Aug	0.85
Sep	0.81
Oct	0.86
Nov	0.69
Dec	0.83

Table 16.2: Pearson correlation coefficient between modelled and observed sea surface oxygen concentration at each grid cell. The correlation was calculated monthly and using the annual mean in each grid cell.

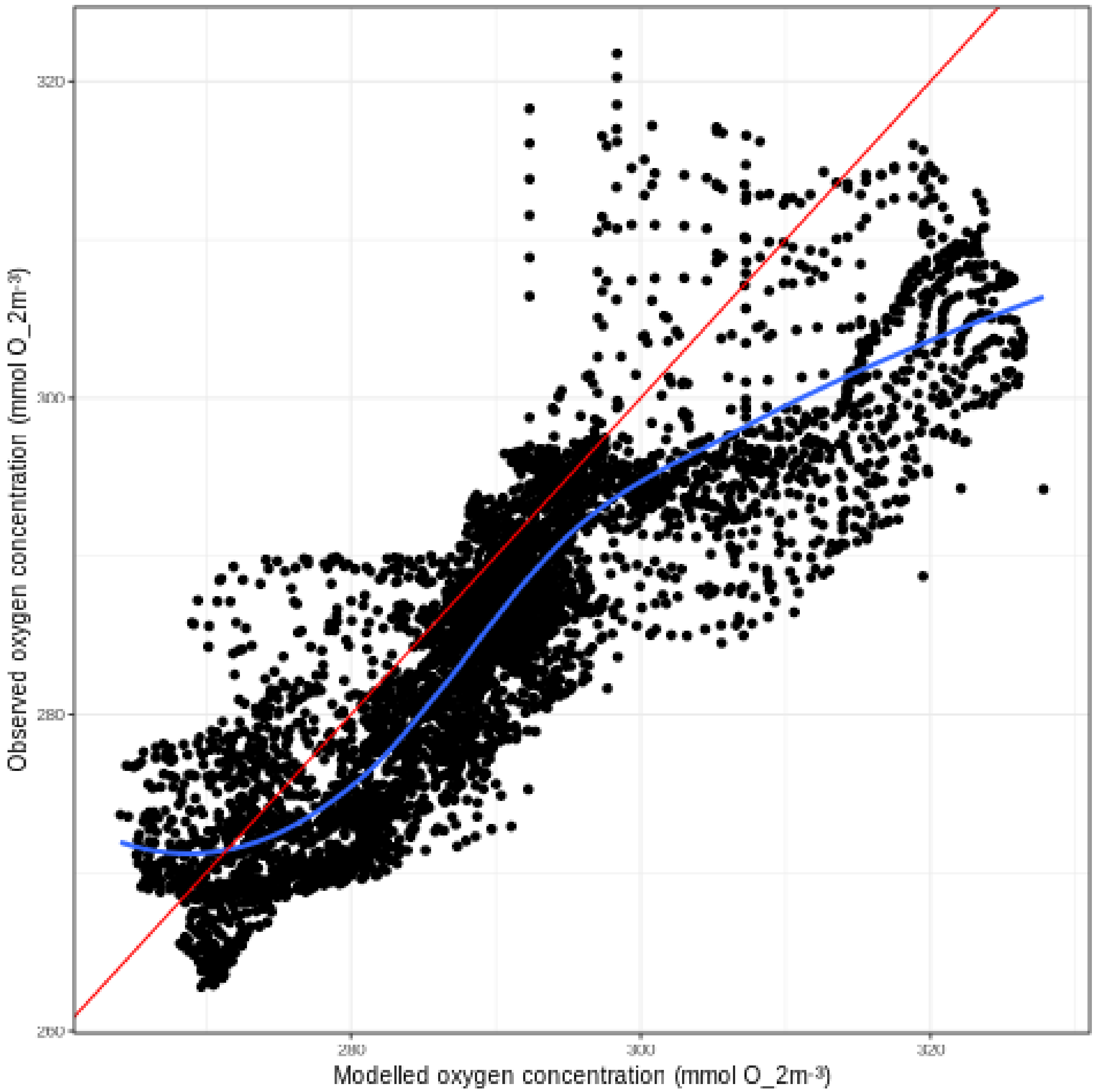


Figure 16.7: Scatter plot of modelled and observed annual average sea surface oxygen concentration in grid cells. The red line is the 1:1 line. The blue line is the linear regression of the modelled and observed values. The slope of the blue line is the slope of the regression line.

16.6 Data Sources for validation of oxygen concentration

Hinrichs, Iris; Gouretski, Viktor; Paetsch, Johannes; Emeis, Kay; Stammer, Detlef (2017). North Sea Biogeochemical Climatology (Version 1.1).

URL: https://www.cen.uni-hamburg.de/en/icdc/data/ocean/nsbc.html

CHAPTER

SEVENTEEN

VALIDATION OF PH USING POINT OBSERVATIONS

17.1 Performance of model sea surface pH

Values from the **top 5m** of the water column were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 3,411 values extracted from the observational database. The map below shows the locations of the matched up data for pH.

The following model output was used to compare with observational values: **O3_pH**.

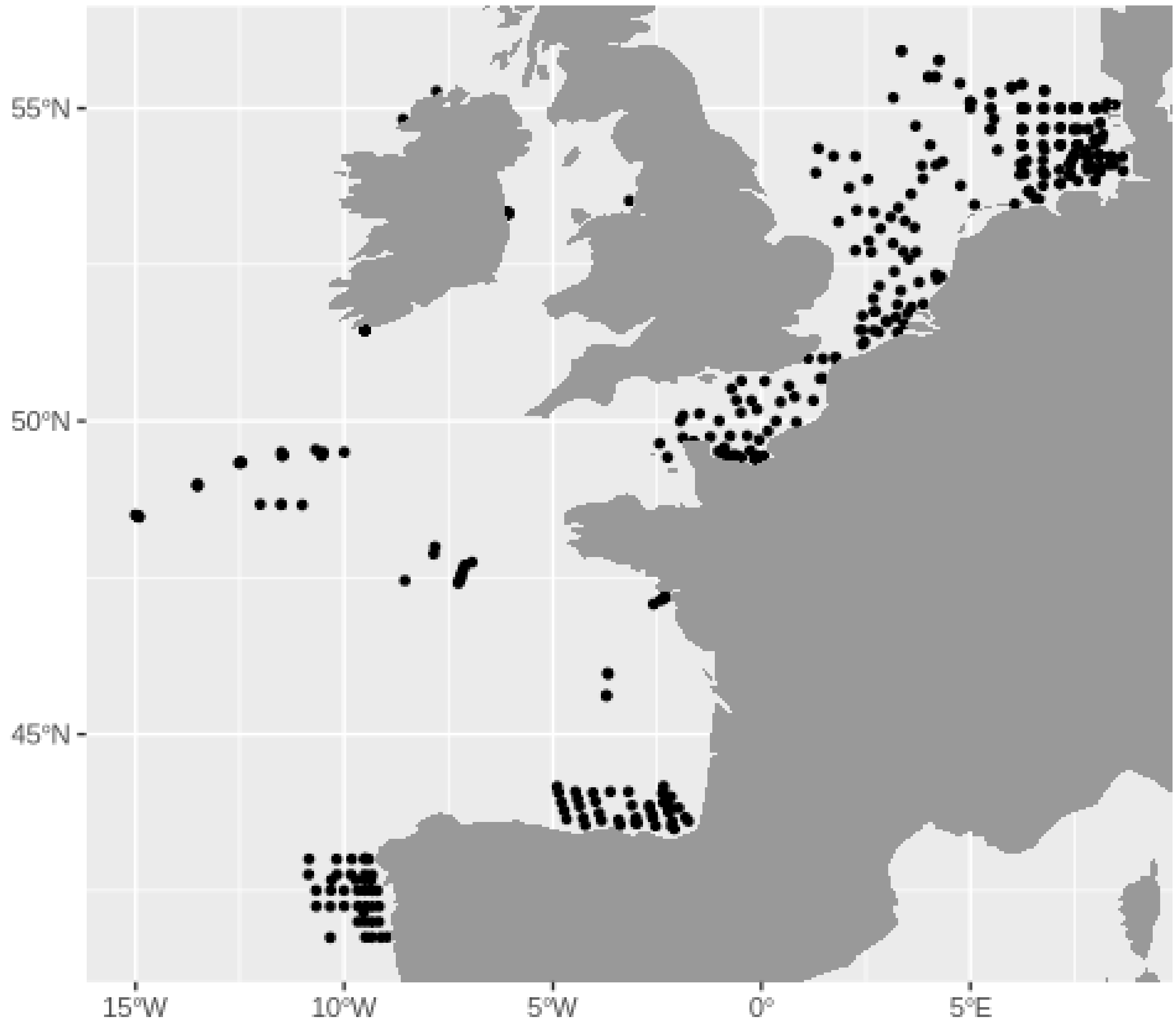


Figure 17.1: Locations of matchups between simulated and observed pH in the top 5m of the water column.

The number of observations in each month ranged from 150 in December to 389 in January. Figure 17.2 below shows the distribution of observations in each month.

Figure 17.2 below shows the bias between the model and observational data for pH. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

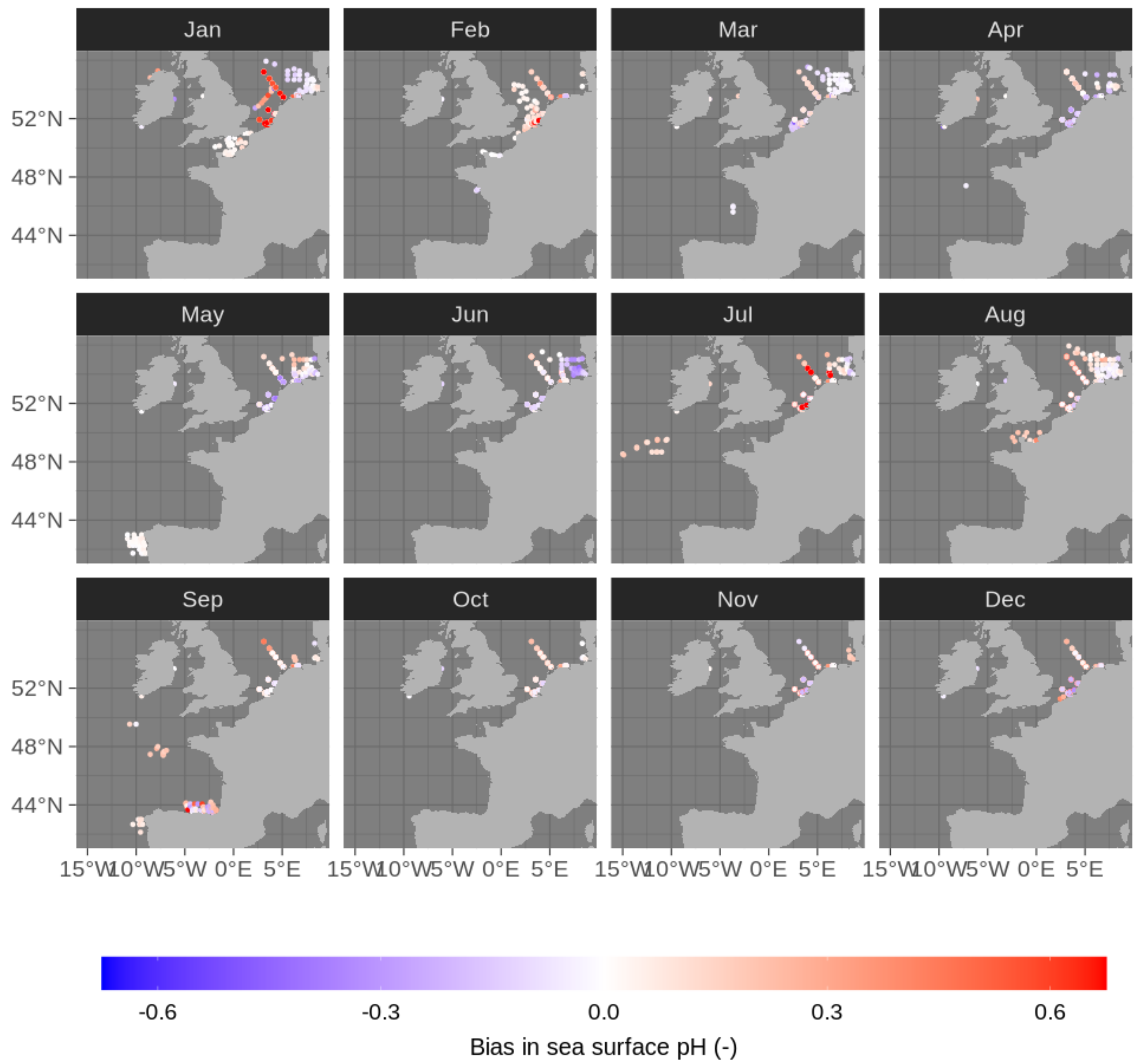


Figure 17.2: Bias in sea surface pH. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 17.3 and 17.4 show the distribution of sea surface pH observations in the model and observational datasets. This is shown for each month of the year (Figure 17.3) and for the entire year (Figure 17.4).

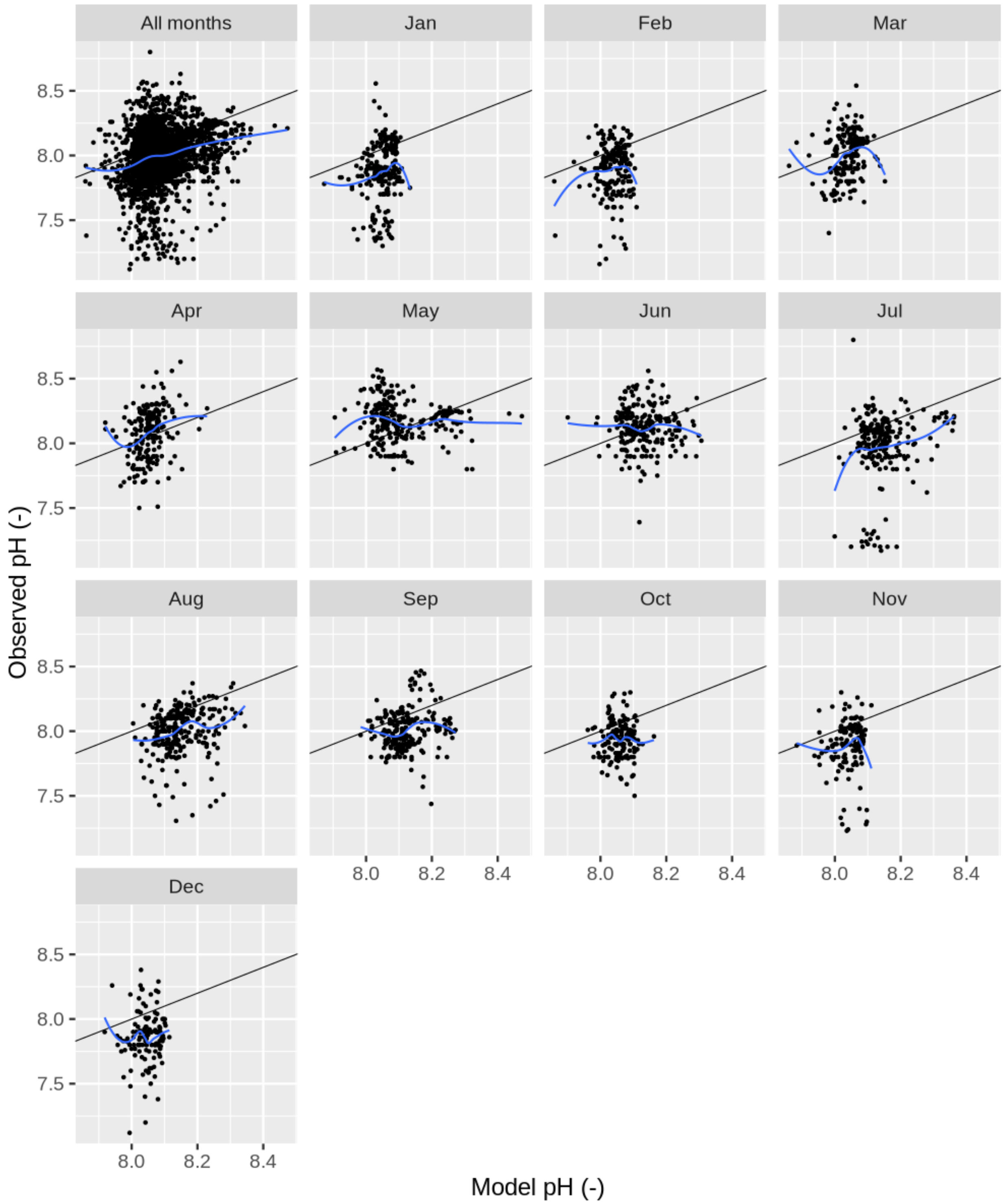


Figure 17.3: Simulated versus observed pH in the top 5m of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

17.2 Summary statistics for sea surface pH

The overall ability of the model to predict the observed pH was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	0.08	0.23	0.20	3,411
Jan	0.14	0.26	0.23	389
Feb	0.11	0.21	0.13	332
Mar	0.04	0.18	0.24	243
Apr	-0.03	0.20	0.30	251
May	-0.09	0.21	-0.07	340
Jun	0.00	0.18	-0.08	336
Jul	0.15	0.28	0.20	337
Aug	0.14	0.24	0.17	358
Sep	0.11	0.22	0.11	367
Oct	0.12	0.19	0.00	157
Nov	0.15	0.25	0.08	151
Dec	0.20	0.28	0.05	150

Table 17.1: Average bias (-) and root-mean square deviation (-) for the model's sea surface pH for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed pH was performed. The modelled pH was used as the independent variable and the observed pH was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.66	2.63	0.05	0.00
Jan	1.22	-1.90	0.04	0.00
Feb	0.52	3.71	0.01	0.16
Mar	1.12	-1.03	0.07	0.00
Apr	1.24	-1.92	0.08	0.00
May	-0.07	8.75	0.00	0.48
Jun	-0.08	8.76	0.00	0.65
Jul	0.76	1.78	0.04	0.00
Aug	0.57	3.39	0.04	0.00
Sep	0.36	5.08	0.02	0.03
Oct	0.02	7.79	0.00	0.95
Nov	0.36	5.02	0.00	0.40
Dec	0.31	5.34	0.00	0.47

Table 17.2: Linear regression analysis of modelled and observed pH. The modelled pH was used as the independent variable and the observed pH was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

17.3 Performance of model near-bottom ph

Near-bottom values of pH were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The near-bottom was defined as observations **within 2m of the seabed**. This was interpolated to the observational grid using the $GEBCO_2$ bathymetry dataset. Model values were interpolated to the observational dataset's longitudes and latitudes using 3D interpolation. **Note:** this analysis has been restricted to observations on the shelf region. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 155 values extracted from the observational database. The map below shows the locations of the matched up data for pH.

The following model output was used to compare with observational values: **O3_pH**.

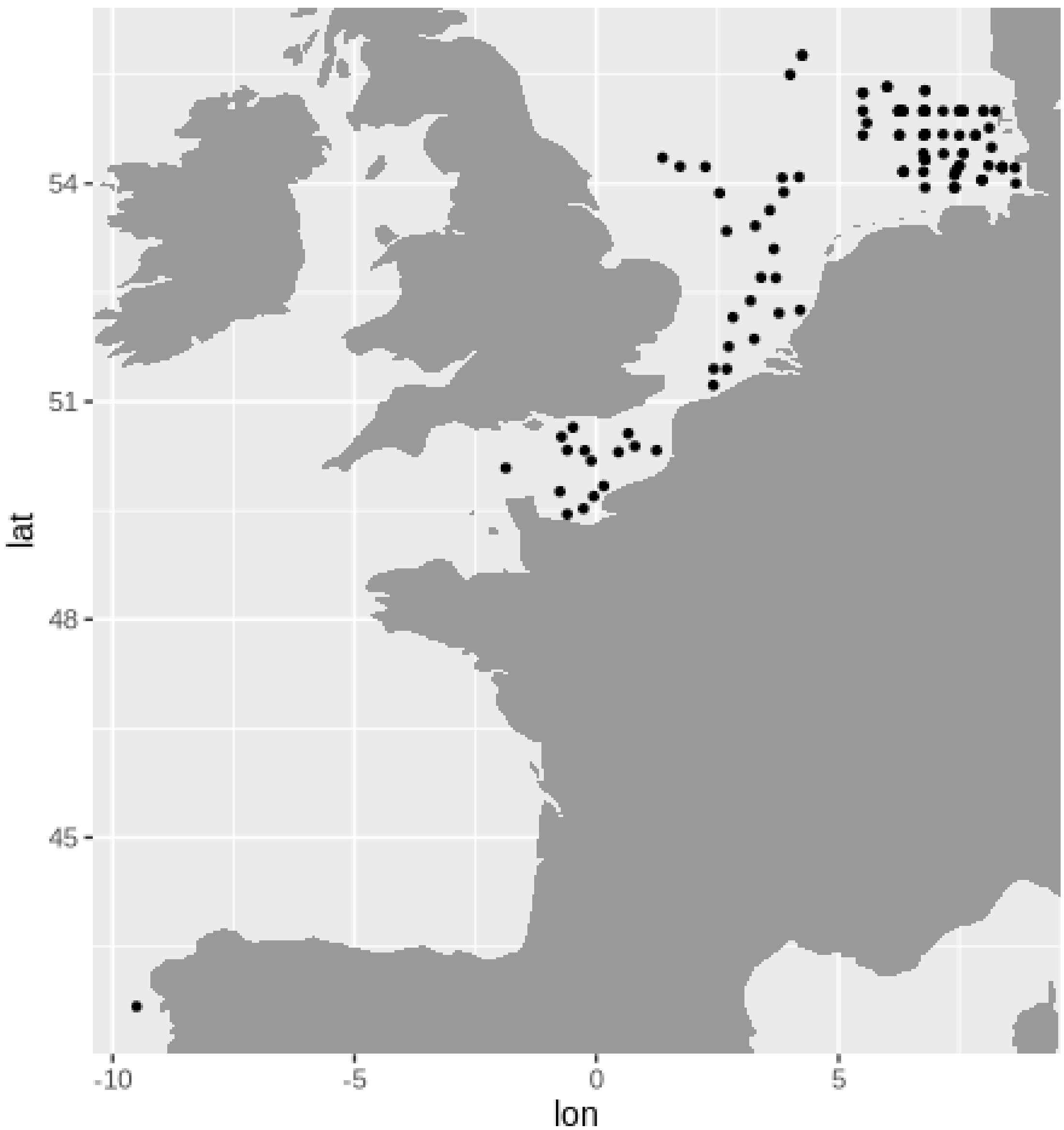


Figure 17.4: Locations of matchups between simulated and observed pH near the bottom of the water column.

The number of observations in each month ranged from 1 in September to 55 in January. Figure 17.5 below shows the distribution of observations in each month.

Figure 17.5 below shows the bias between the model and observational data for pH. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

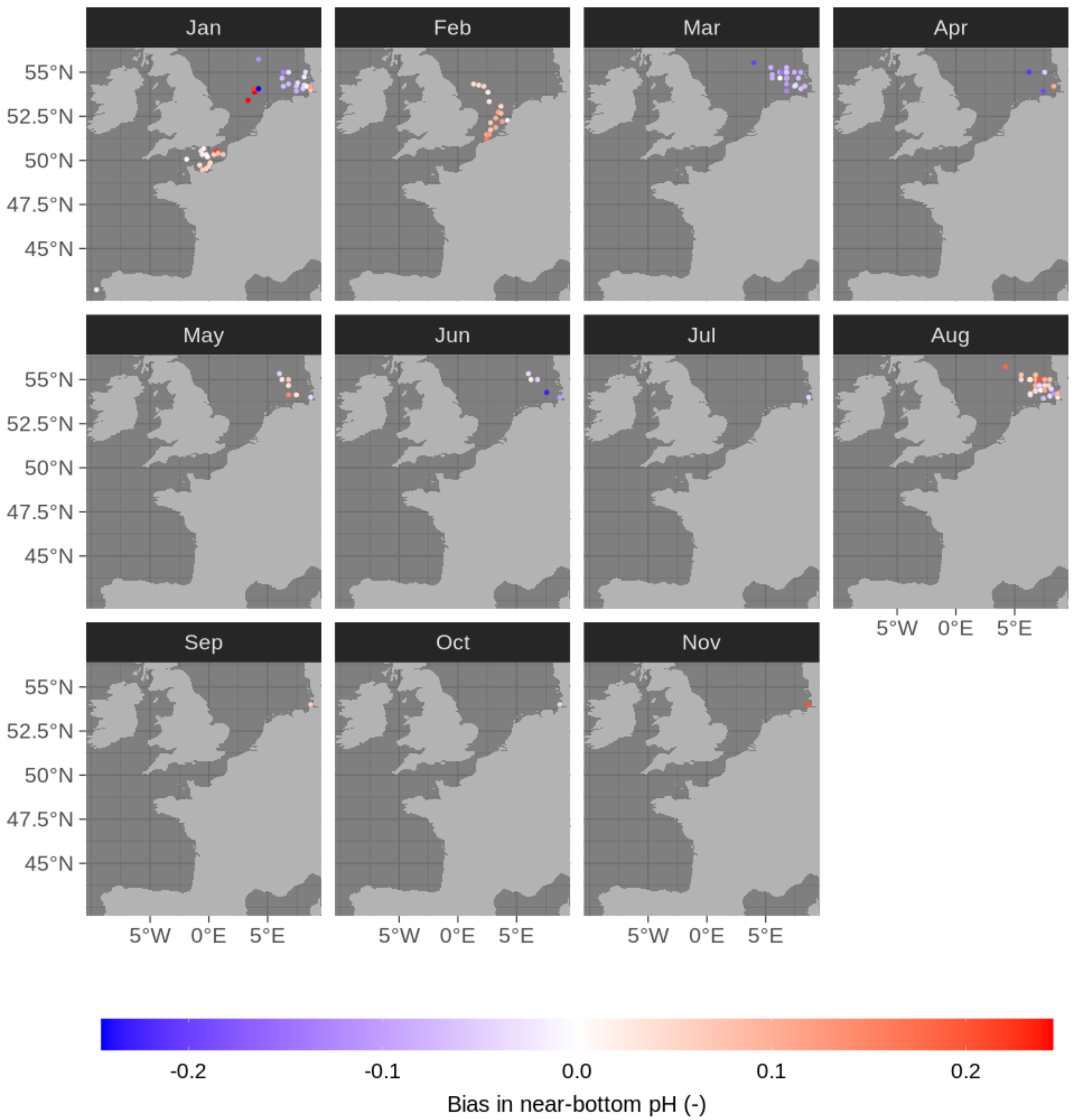


Figure 17.5: Bias in near-bottom pH. The bias is calculated as model-observation. The colour scale is from blue (negative

bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 17.6 and 17.7 show the distribution of near-bottom pH observations in the model and observational datasets. This is shown for each month of the year (Figure 17.6) and for the entire year (Figure 17.7).

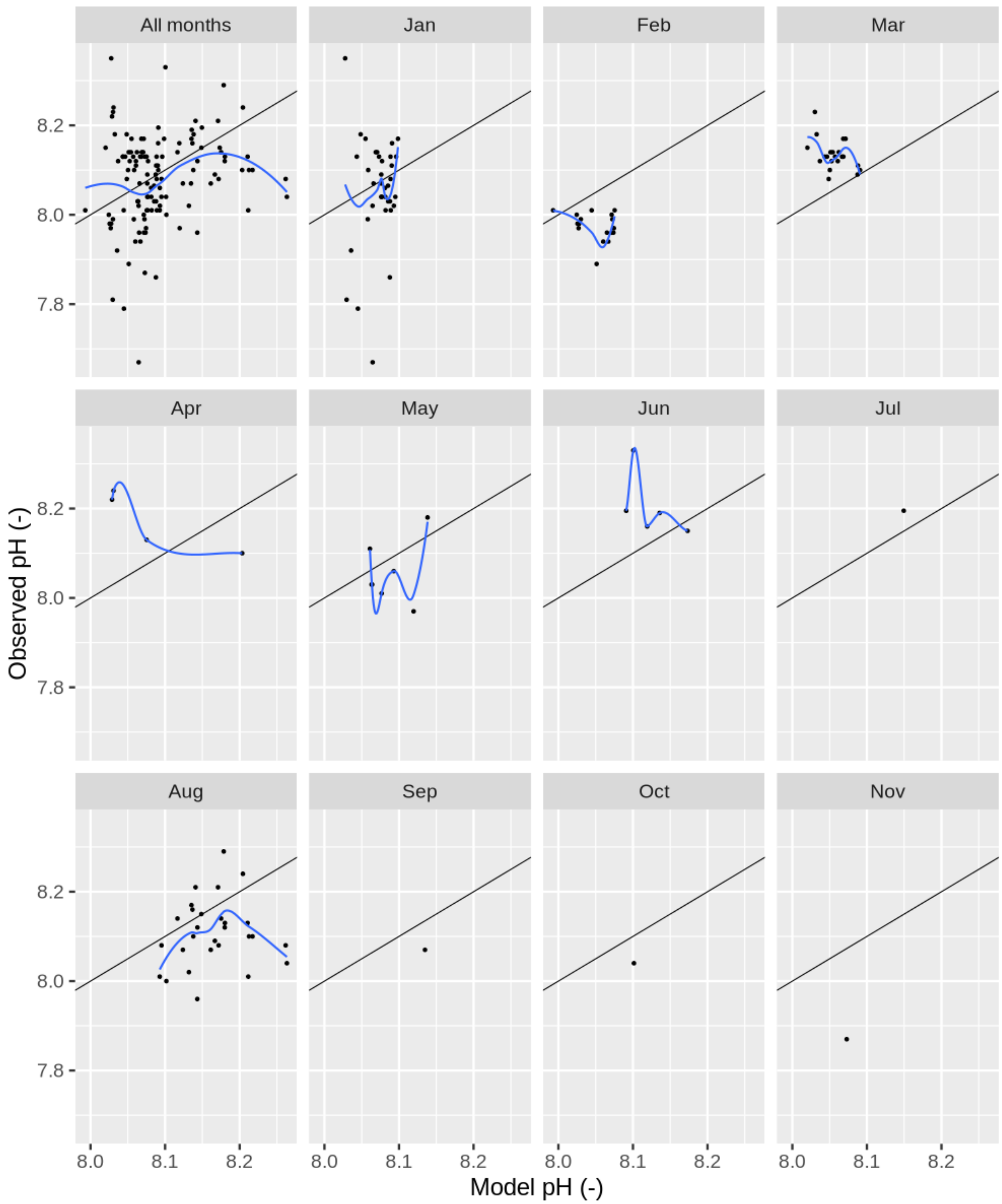


Figure 17.6: Simulated versus observed pH near the bottom of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

17.4 Summary statistics for near-bottom pH

The overall ability of the model to predict the observed pH was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	0.03	0.13	0.24	155
Jan	0.05	0.18	0.11	55
Feb	0.08	0.09	-0.25	26
Mar	-0.08	0.09	-0.44	21
Apr	-0.09	0.15	-0.87	4
May	0.02	0.07	0.48	8
Jun	-0.08	0.12	0.06	6
Jul	-0.05	0.06	1.00	2
Aug	0.05	0.10	0.14	30
Sep	0.06	0.06	N/A	1
Oct	0.06	0.06	N/A	1
Nov	0.20	0.20	N/A	1

Table 17.3: Average bias (-) and root-mean square deviation (-) for the model's near-bottom pH for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed pH was performed. The modelled pH was used as the independent variable and the observed pH was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.39	4.93	0.04	0.03
Jan	0.71	2.29	0.01	0.52
Feb	-0.36	10.90	0.08	0.27
Mar	-0.79	14.49	0.20	0.04
Apr	-0.72	13.97	0.75	0.13
May	0.69	2.47	0.09	0.51
Jun	-1.24	18.32	0.31	0.33
Aug	0.20	6.48	0.01	0.55

Table 17.4: Linear regression analysis of modelled and observed pH. The modelled pH was used as the independent variable and the observed pH was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

17.5 Depth-resolved comparisons of modelled and observed pH

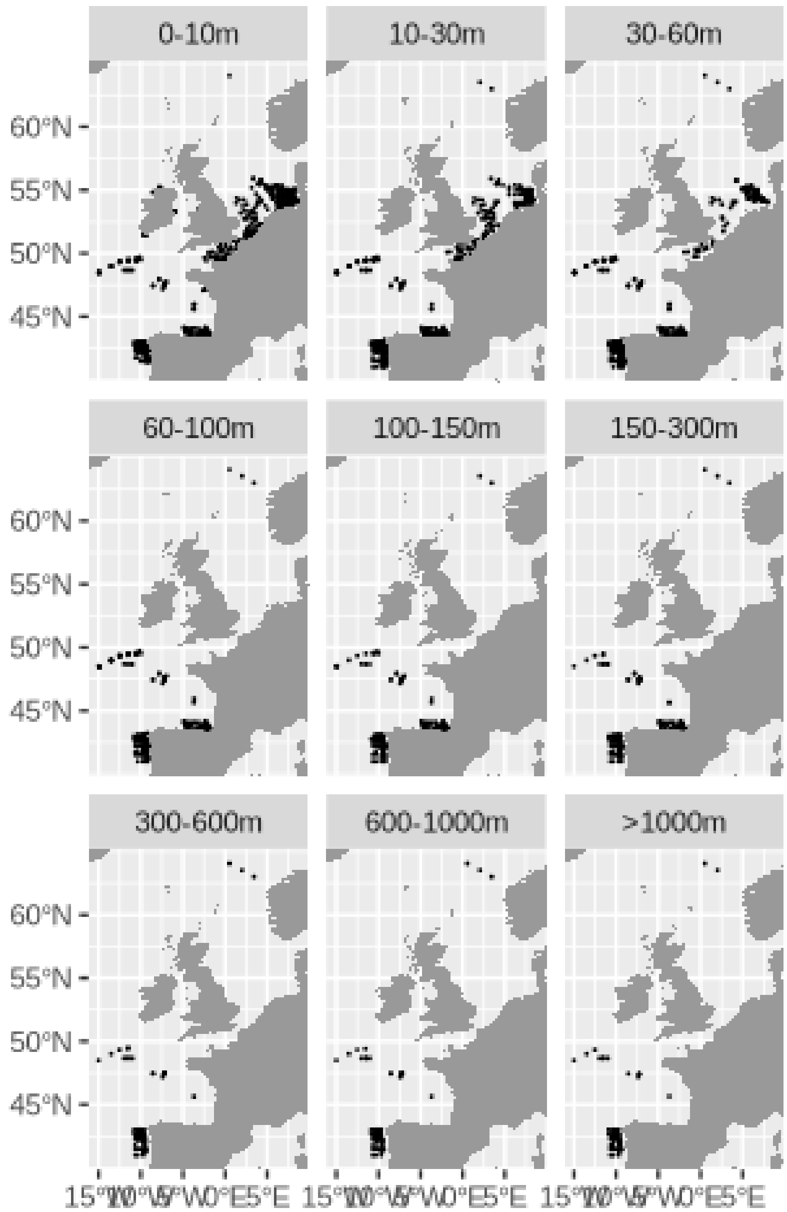


Figure 17.7: The geographic distribution of matchups between the model and observational pH. The data has been binned into depth ranges. The depth ranges are 0-10m, 10-30m, 30-60m, 60-100m, 100-150m, 150-300m, 300-600m, 600-1000m, and >1000m. The number of observations in each depth range is shown in the tables below.

Depth	Bias	*r*	RMSD	Number of observations
0-150m	0.01	0.21	0.20	9,334
0-10m	0.06	0.25	0.18	1,496
10-30m	0.07	0.32	0.20	2,163
30-60m	-0.00	0.24	0.21	1,842
60-100m	-0.04	0.29	0.22	1,858
100-150m	-0.02	0.18	0.21	1,975
150-300m	0.14	0.47	0.22	884
300-600m	0.01	0.10	0.09	202
600-1000m	-0.00	-0.33	0.08	235
>1000m	0.00	-0.07	0.04	235

Table 17.5: Average bias (-), root-mean square difference (RMSD) and correlation coefficient of modelled and observed pH for different depth ranges. The bias is calculated as model-observation. The RMSD is the square root of the mean squared difference. The correlation coefficient is the Pearson correlation coefficient between the model and observed values.

17.6 Data Sources for validation of pH

ICES Data Portal, Dataset on Ocean HydroChemistry, Extracted March 3, 2023. ICES, Copenhagen

CHAPTER
EIGHTEEN

SEA SURFACE PHOSPHATE VALIDATION USING GRIDDED OBSERVATIONS FROM NSBC

We used version 1.1 of the **North Sea Biogeochemical Climatology** (NSBC) to validate **sea surface phosphate**. NSBC is a monthly climatology that covers the region 47°-65°N and 15°W-15°E. The data is made up of observations over the period 1960-2014. For validation purposes we only used the level 3 data, which a gridded monthly climatology at a spatial resolution of 1/4°. The data can be download from NSBC.

Matchup procedure: The model and observations were matched up as follows. First, the model dataset was cropped by a small amount to make sure cells close to the boundary were removed. The model was then regridded to the observational grid if the observational grid was coarser using nearest neighbour. Only grid cells with model and observational data were maintained. The following model output was used to compare with the observational values: **N1_p**.

18.1 Baseline climatologies of sea surface phosphate

Climatologies of model and observational sea surface phosphate concentration are shown in the figures below. The model climatology is calculated using the years **1986-2017**.

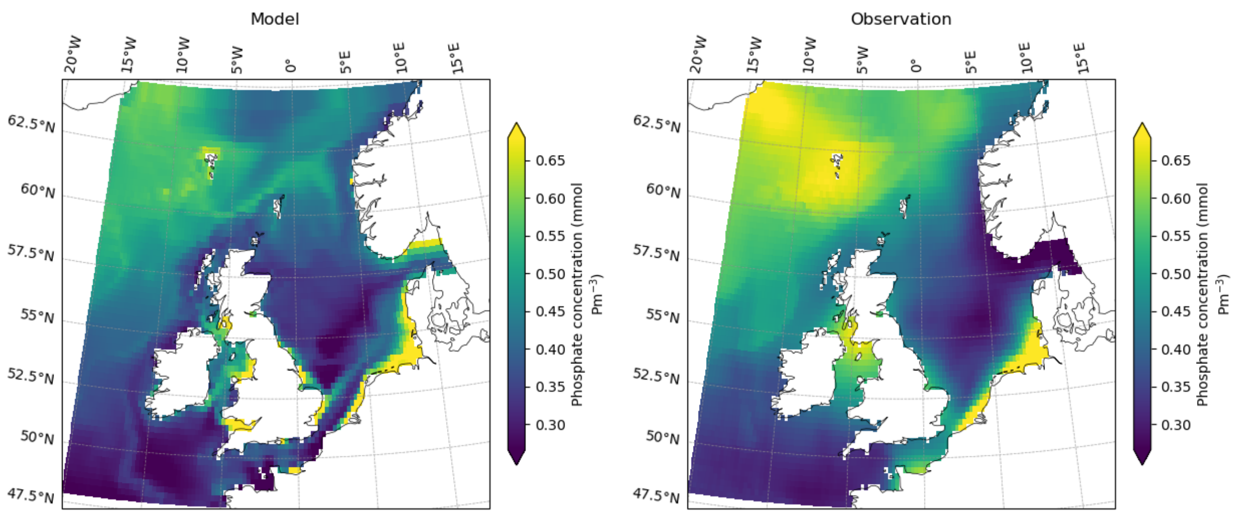


Figure 18.1: Annual average surface phosphate concentration from the model (1986-2017) and observations. Data is limited to the 2nd and 98th percentile of the combined model and observational data. Arrows indicate that values can exceed the colourbar limits.

18.2 Assessing model bias for surface phosphate concentration

Figure 18.2 shows the average bias of surface phosphate concentration simulated by the model. A positive bias indicates that the model overestimates the observation, while a negative bias indicates that the model overpredicts the observation.

The spatial average bias of surface phosphate concentration is -0.04 mmol Pm^{-3}. Overall, the model underestimates the observations in 81.4% of the model domain.

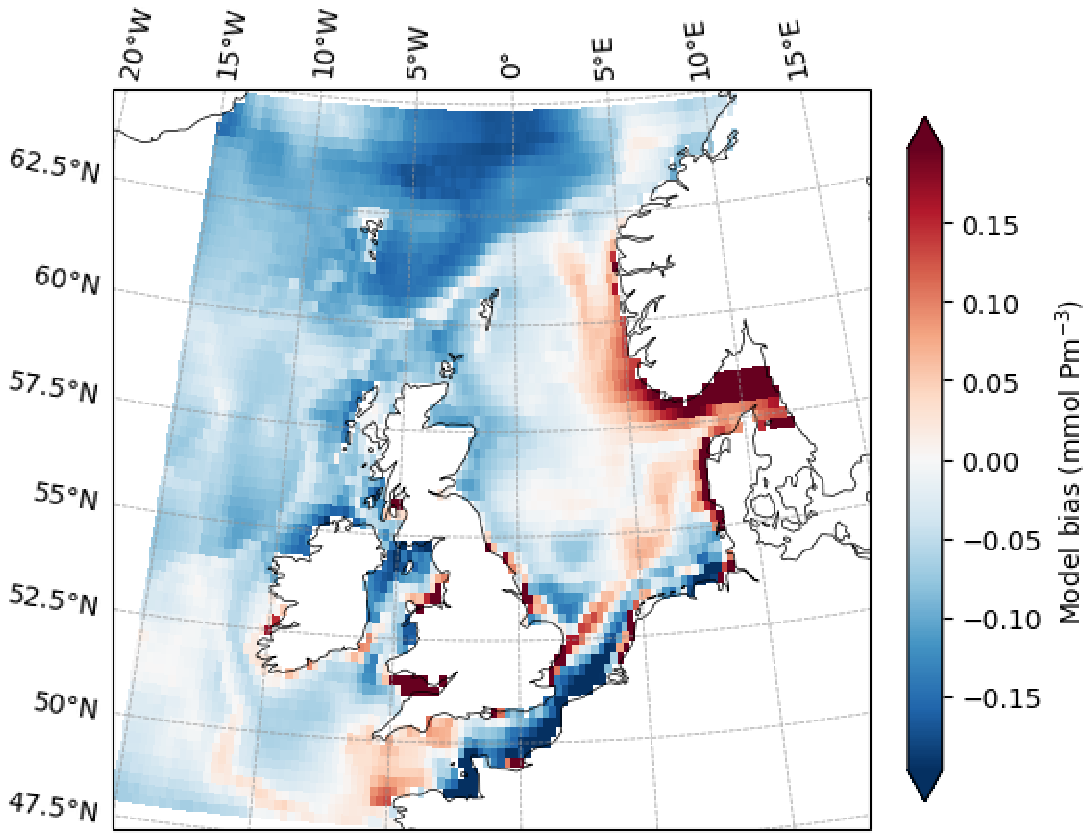


Figure 18.2: Bias of surface phosphate concentration from the model. A positive bias indicates that the model overestimates the observation. For clarity, the colourbar is limited to the 2nd and 98th percentile of the data.

18.3 Can the model reproduce seasonality of sea surface phosphate concentration?

The ability of the model to reproduce seasonality of sea surface phosphate concentration was assessed by comparing the modelled and observed seasonal cycle of phosphate concentration. First, we derive a monthly climatology for the model data. Then, we calculate the Pearson correlation coefficient between the modelled and observed phosphate concentration at each grid cell.

Note: we are only assessing the ability of the model to reproduce the ability of the model to reproduce seasonal changes, not long-term trends.

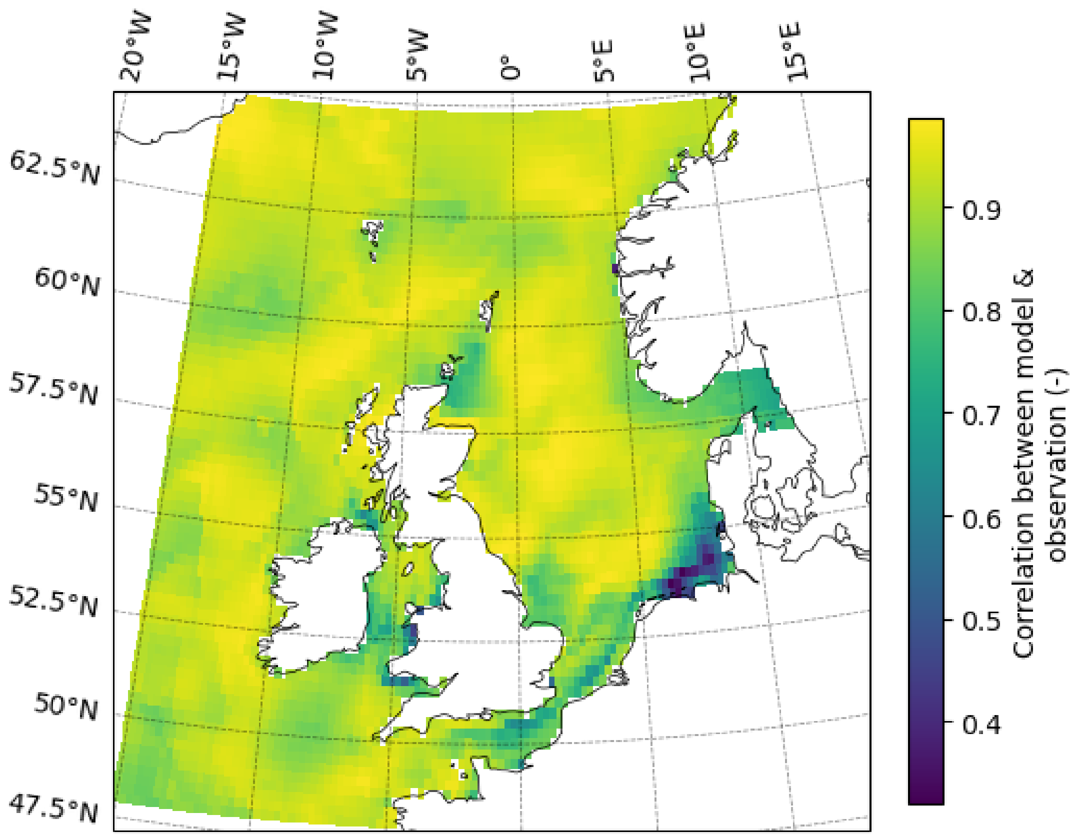
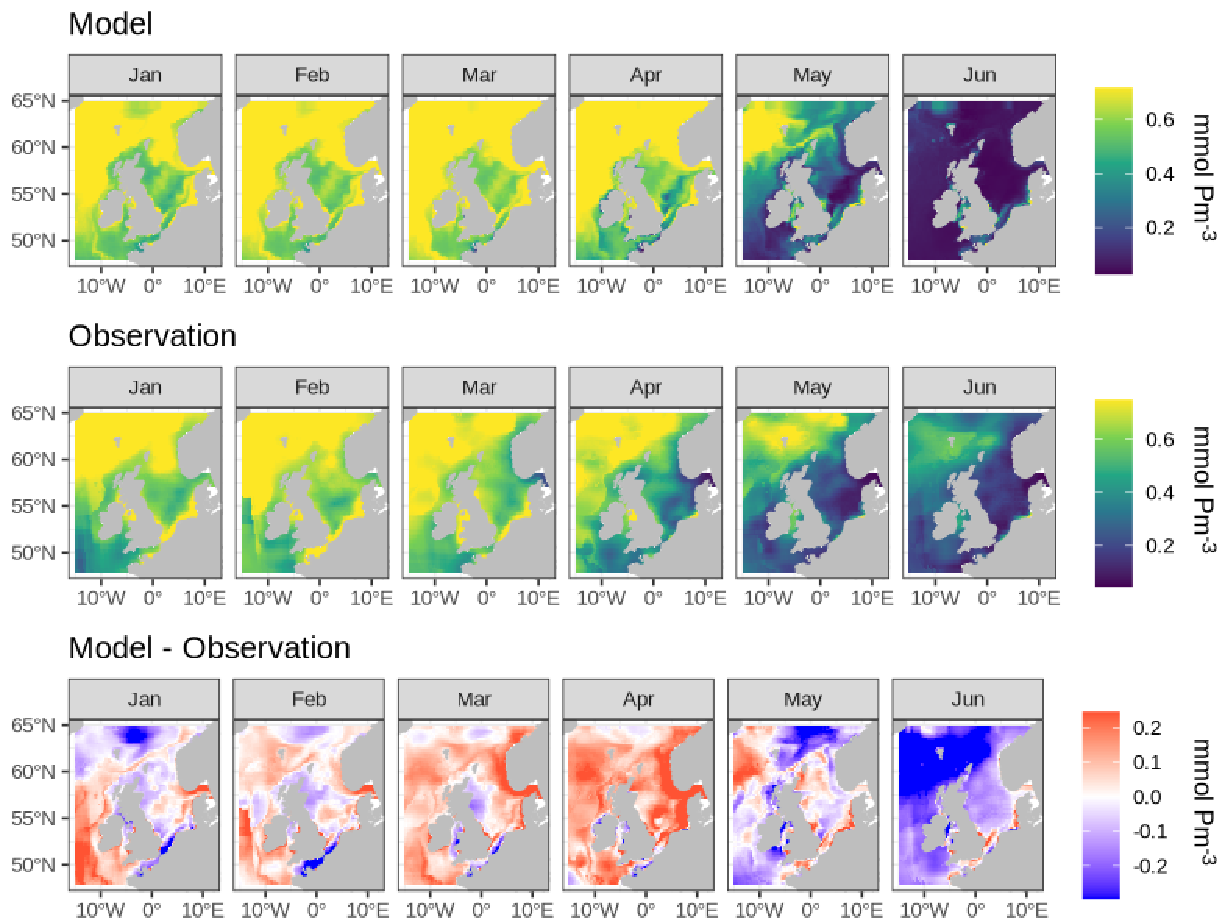


Figure 18.3: Seasonal temporal correlation between model and observations for surface phosphate concentration. This is the Pearson correlation coefficient between climatology monthly mean values in the model and observations.

The seasonal cycles of simulated and observed phosphate concentration are compared in Figure 18.4 below. This figure shows the model and observation average in each month of the year, and the differences between the two each month


Model
Jan
Feb
Mar
Apr
May
Jun
65°N
60°N
55°N
50°N
10°W 0° 10°E
0.6
0.4
0.2
mmol Pm-3
Observation
Model - Observation
0.2
0.1
0.0
-0.1
-0.2

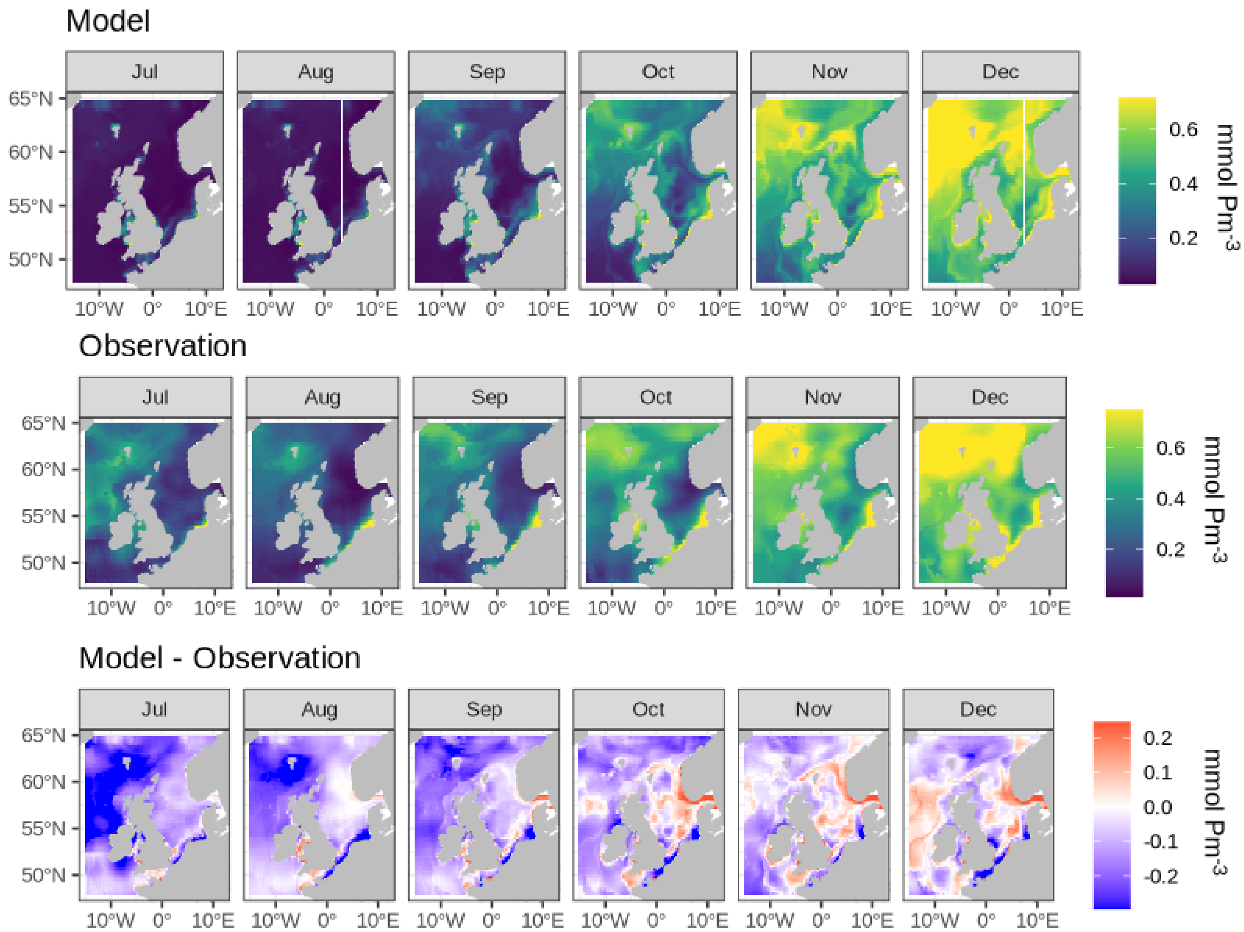


Figure 18.4: Monthly mean surface phosphate concentration for the model, observation and the difference between model and observations. For clarity, the maximum values are capped to the 98th percentiles.

18.4 Regional assessment of model performance for sea surface phosphate concentration

We assessed the regional performance of the model by comparing the model with observations in a number of regions. The regions considered are mapped below.

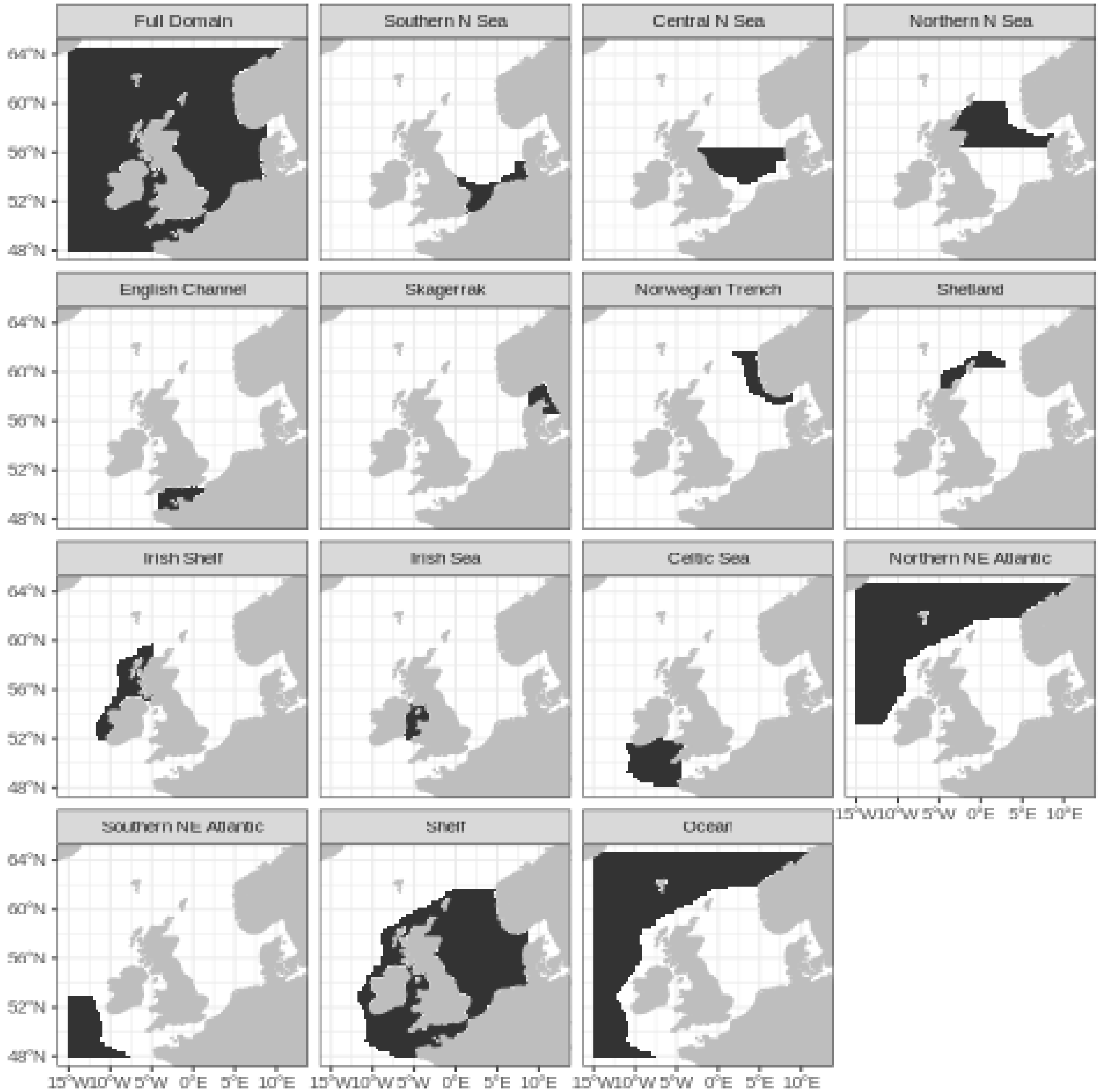


Figure 18.5: Regions used for validation of sea surface phosphate concentration.

Time series were constructed comparing the monthly mean of the spatial average sea surface phosphate concentration in each region. The spatial average was calculated using the mean of all grid cells within each region, accounting for grid cell area.

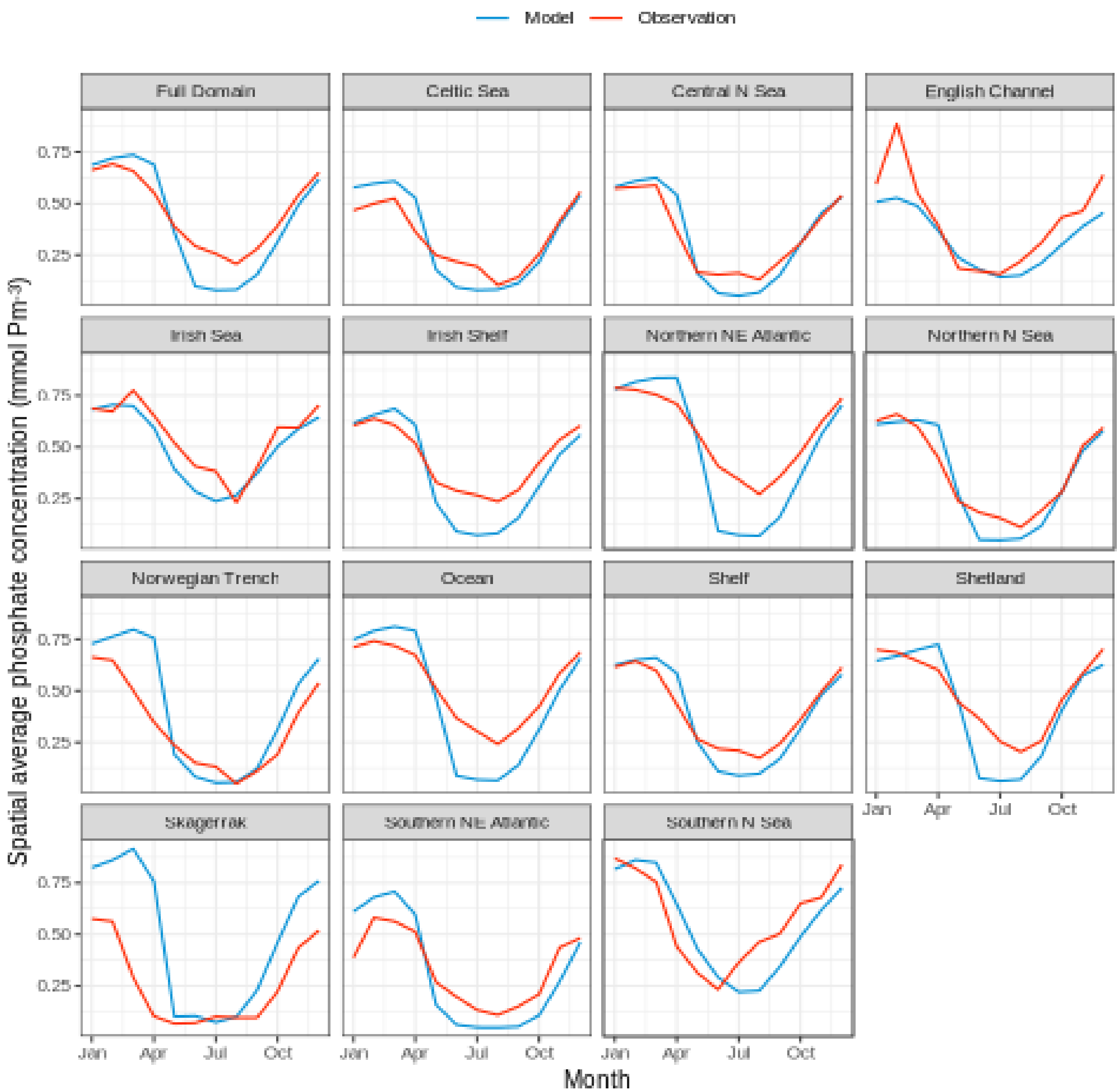


Figure 18.6: Seasonal cycle of sea surface phosphate concentration for model and observations for each region. The spatial average is taken over the region.

The table below shows the average bias of sea surface phosphate concentration in each region. The bias is calculated as the modelled value minus the observed value. A positive bias indicates that the model overestimates the observed value,

while a negative bias indicates that the model underestimates the observed value.

Region	Temporal correlation	Bias	RMSD
Full Domain	0.91	-0.05	0.17
Celtic Sea	0.91	0.00	0.23
Central North Sea	0.93	-0.01	0.11
Channel	0.87	-0.09	0.20
Irish Sea	0.84	-0.05	0.21
Irish Shelf	0.92	-0.07	0.14
Northern North East Atlantic	0.93	-0.08	0.18
Northern North Sea	0.93	-0.02	0.10
Norwegian Trench	0.89	0.09	0.19
Ocean	0.92	-0.07	0.17
Shelf	0.90	-0.02	0.18
Shetland	0.92	-0.06	0.13
Skagerrak	0.77	0.23	0.34
Southern North East Atlantic	0.91	-0.02	0.13
Southern North Sea	0.75	-0.03	0.30

Table 18.1: Summary of performance of the model sea surface phosphate concentration in each region. The bias (mmol Pm-3) column represents the spatial average of the annual mean modelled value minus the observed value. The temporal correlation column represents the spatial mean of the temporal correlation between the model and observations per grid cell.

18.5 Can the model reproduce spatial patterns of sea surface phosphate concentration?

The ability of the model to reproduce spatial patterns of sea surface phosphate concentration was assessed by comparing the modelled and observed phosphate concentration at each grid cell. We calculated the Pearson correlation coefficient between the modelled and observed phosphate concentration at each grid cell.

This was carried out monthly and using the annual mean in each grid cell

Time period	*r*
Annual mean	0.48
Jan	0.46
Feb	0.39
Mar	0.47
Apr	0.51
May	0.66
Jun	0.07
Jul	0.17
Aug	0.25
Sep	0.43
Oct	0.53
Nov	0.51
Dec	0.50

Table 18.2: Pearson correlation coefficient between modelled and observed sea surface phosphate concentration at each grid cell. The correlation was calculated monthly and using the annual mean in each grid cell.

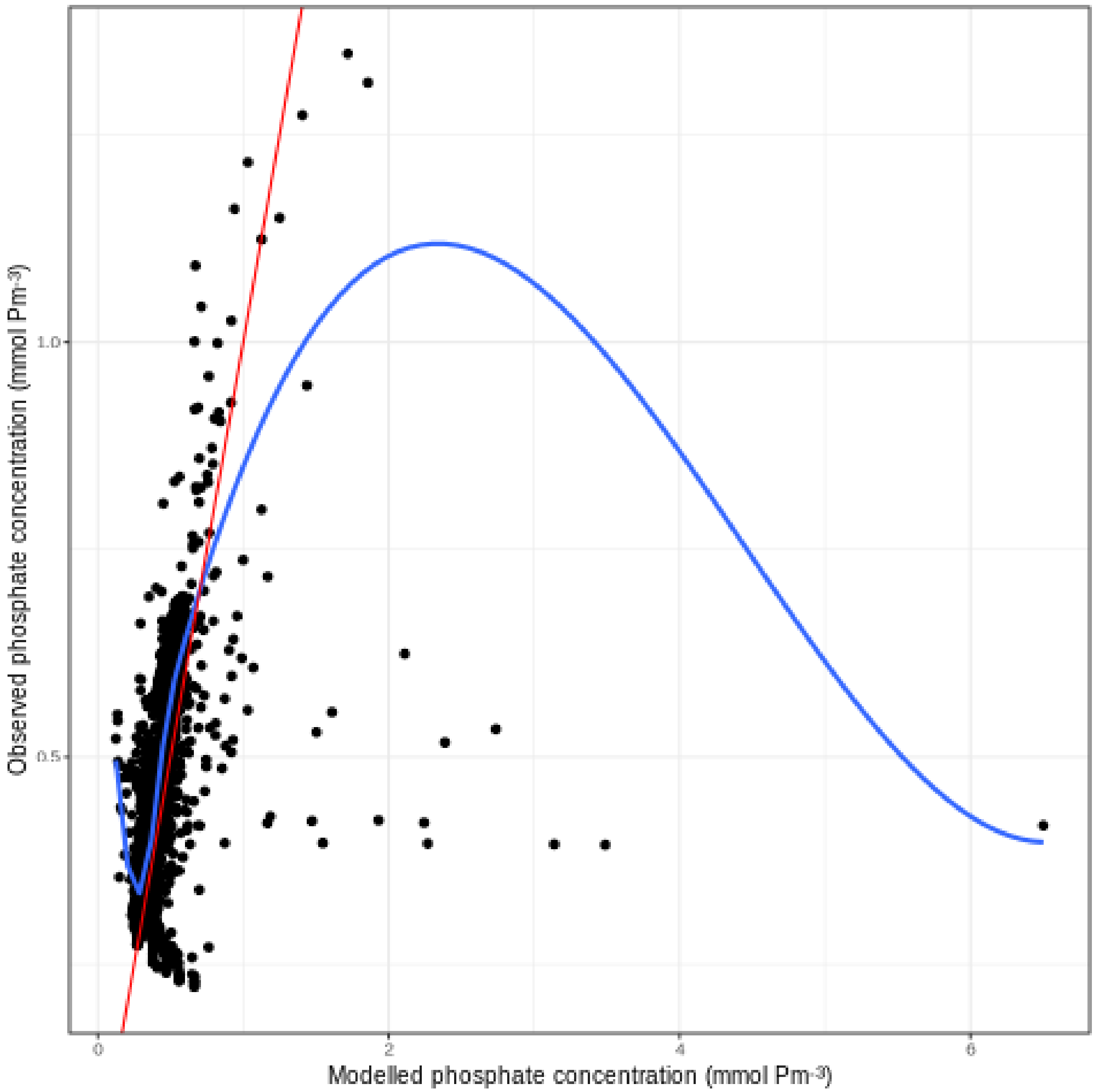


Figure 18.7: Scatter plot of modelled and observed annual average sea surface phosphate concentration in grid cells. The red line is the 1:1 line. The blue line is the linear regression of the modelled and observed values. The slope of the blue line is the slope of the regression line.

18.6 Data Sources for validation of phosphate concentration

Hinrichs, Iris; Gouretski, Viktor; Paetsch, Johannes; Emeis, Kay; Stammer, Detlef (2017). North Sea Biogeochemical Climatology (Version 1.1).

URL: https://www.cen.uni-hamburg.de/en/icdc/data/ocean/nsbc.html

CHAPTER
NINETEEN

SEA SURFACE SALINITY VALIDATION USING GRIDDED OBSERVATIONS FROM NSBC

We used version 1.1 of the **North Sea Biogeochemical Climatology** (NSBC) to validate **sea surface salinity**. NSBC is a monthly climatology that covers the region 47°-65°N and 15°W-15°E. The data is made up of observations over the period 1960-2014. For validation purposes we only used the level 3 data, which a gridded monthly climatology at a spatial resolution of 1/4°. The data can be download from NSBC.

Matchup procedure: The model and observations were matched up as follows. First, the model dataset was cropped by a small amount to make sure cells close to the boundary were removed. The model was then regridded to the observational grid if the observational grid was coarser using nearest neighbour. Only grid cells with model and observational data were maintained. The following model output was used to compare with the observational values: **vosaline**.

19.1 Baseline climatologies of sea surface salinity

Climatologies of model and observational sea surface salinity are shown in the figures below. The model climatology is calculated using the years **1986-2017**.

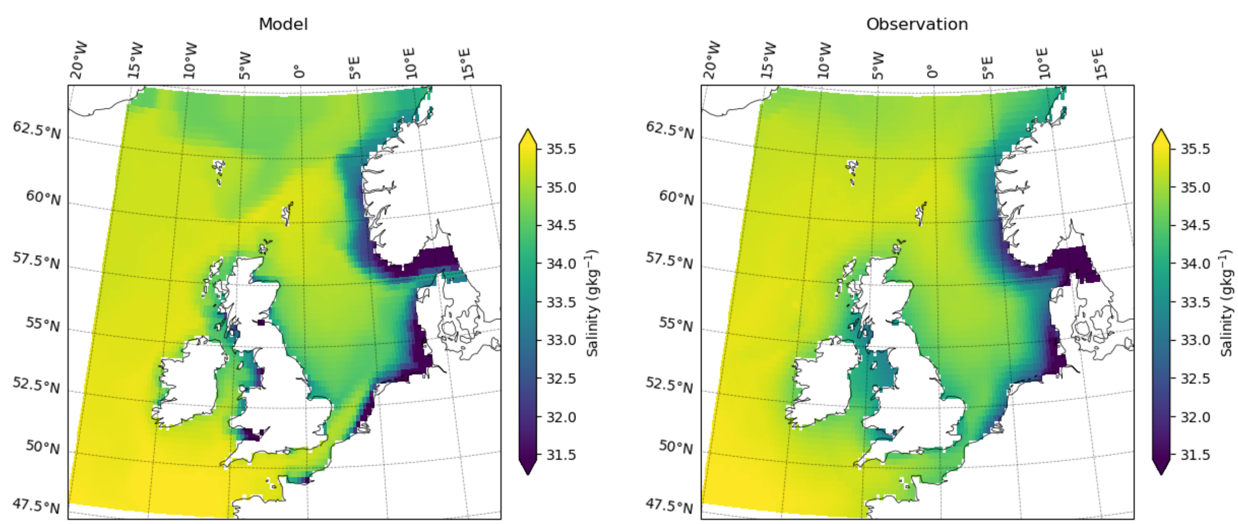


Figure 19.1: Annual average surface salinity from the model (1986-2017) and observations. Data is limited to the 2nd and 98th percentile of the combined model and observational data. Arrows indicate that values can exceed the colourbar limits.

19.2 Assessing model bias for surface salinity

Figure 19.2 shows the average bias of surface salinity simulated by the model. A positive bias indicates that the model overestimates the observation, while a negative bias indicates that the model overpredicts the observation.

The spatial average bias of surface salinity is 0.02 g kg^{-1}. Overall, the model overestimates the observations in 52.4% of the model domain.

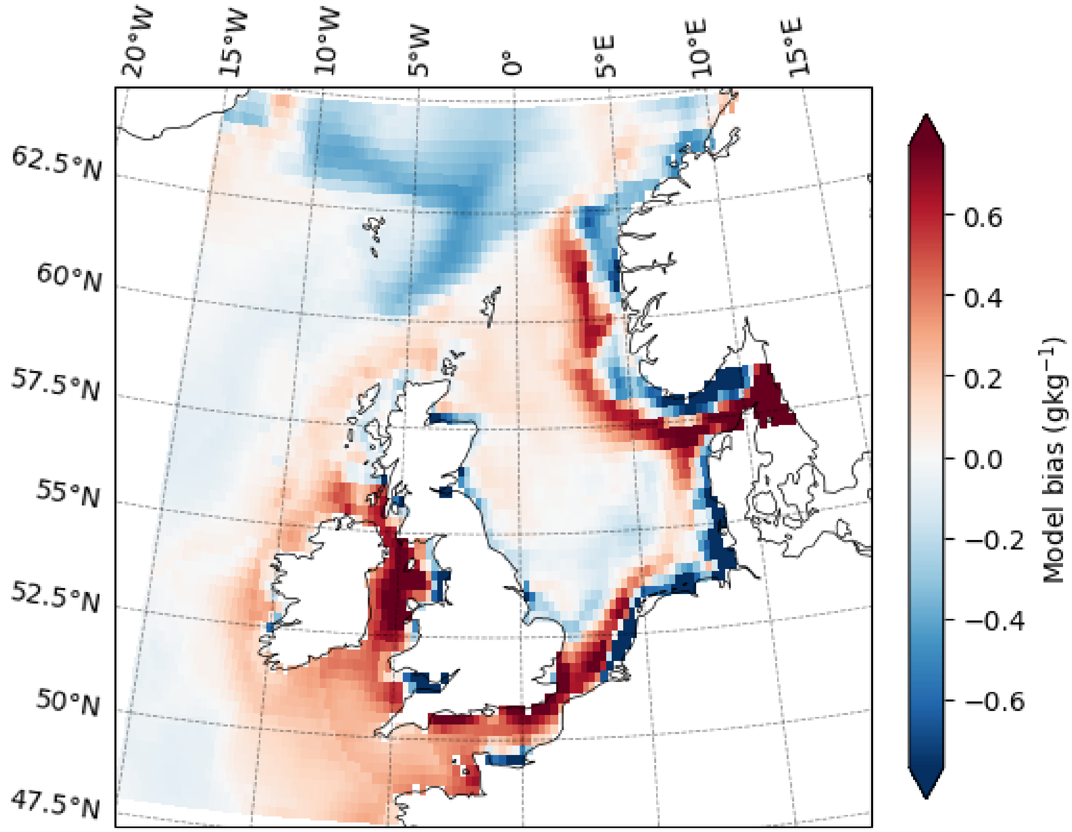


Figure 19.2: Bias of surface salinity from the model. A positive bias indicates that the model overestimates the observation. For clarity, the colourbar is limited to the 2nd and 98th percentile of the data.

19.3 Can the model reproduce seasonality of sea surface salinity?

The ability of the model to reproduce seasonality of sea surface salinity was assessed by comparing the modelled and observed seasonal cycle of salinity. First, we derive a monthly climatology for the model data. Then, we calculate the Pearson correlation coefficient between the modelled and observed salinity at each grid cell.

Note: we are only assessing the ability of the model to reproduce the ability of the model to reproduce seasonal changes, not long-term trends.

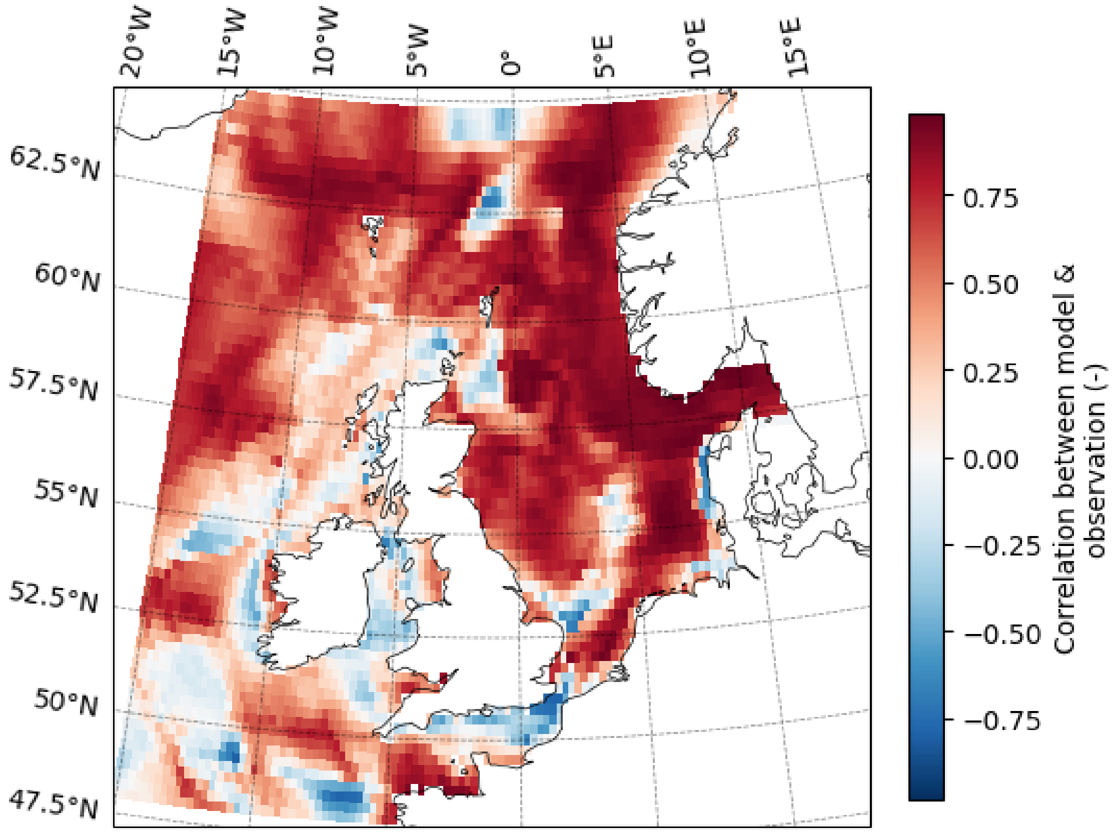
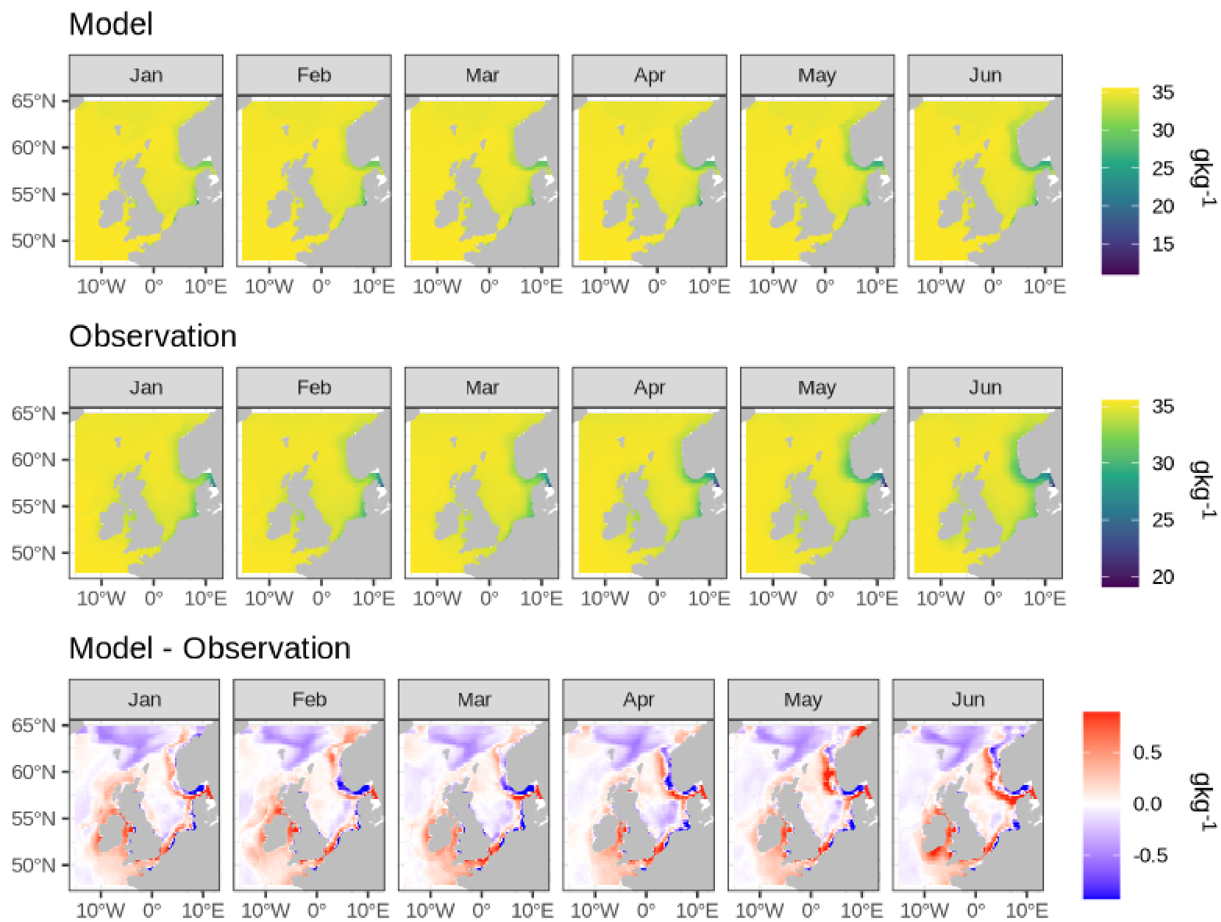


Figure 19.3: Seasonal temporal correlation between model and observations for surface salinity. This is the Pearson correlation coefficient between climatology monthly mean values in the model and observations.

The seasonal cycles of simulated and observed salinity are compared in Figure 19.4 below. This figure shows the model and observation average in each month of the year, and the differences between the two each month



Model
Jan
Feb
Mar
Apr
May
Jun
65°N
60°N
55°N
50°N
10°W 0° 10°E
35
30
25
20
15
gkg-1
Observation
Jan
Feb
Mar
Apr
May
Jun
35
30
25
20
gkg-1
Model - Observation
Jan
Feb
Mar
Apr
May
Jun
0.5
0.0
-0.5
gkg-1

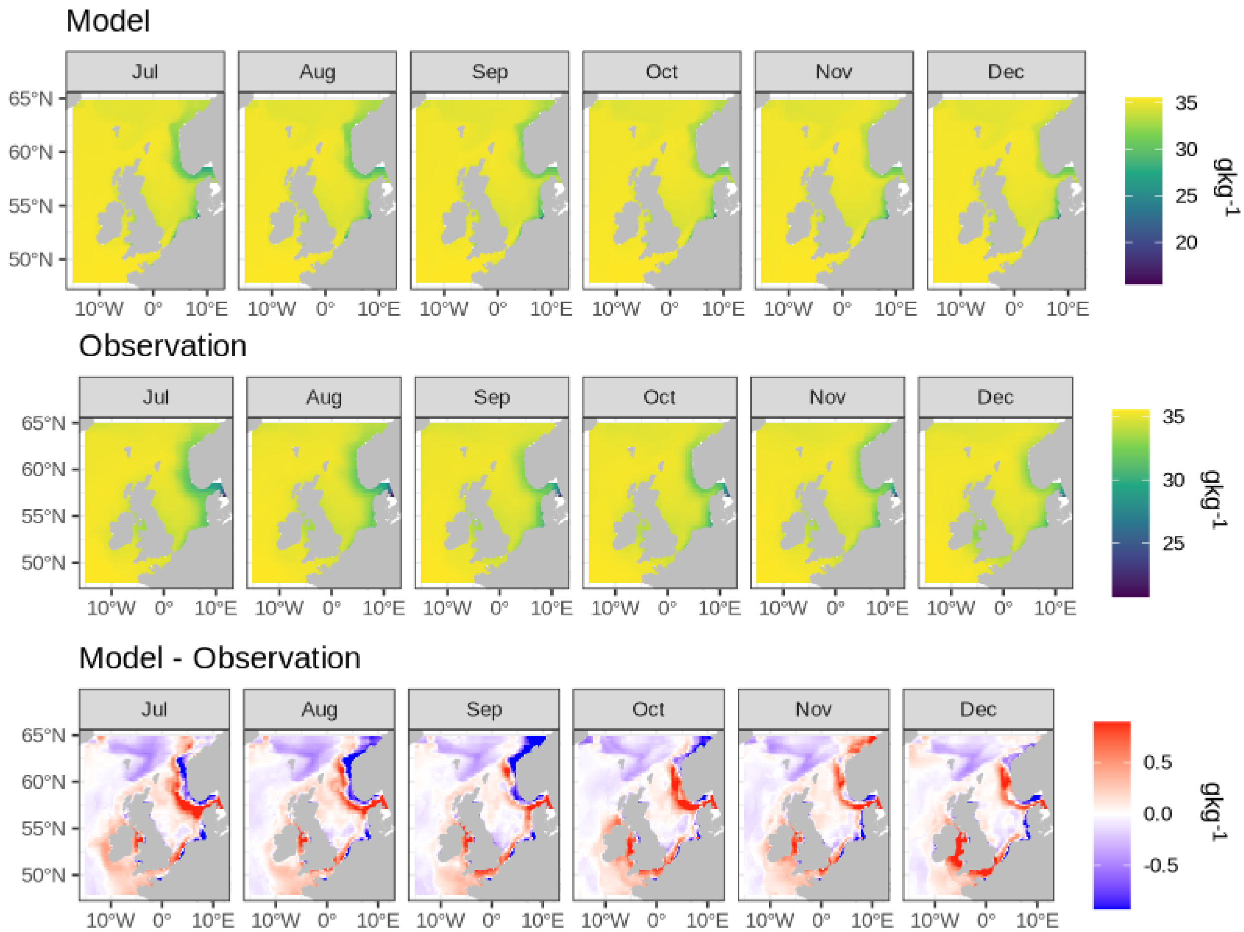


Figure 19.4: Monthly mean surface salinity for the model, observation and the difference between model and observations. For clarity, the maximum values are capped to the 98th percentiles.

19.4 Regional assessment of model performance for sea surface salinity

We assessed the regional performance of the model by comparing the model with observations in a number of regions. The regions considered are mapped below.

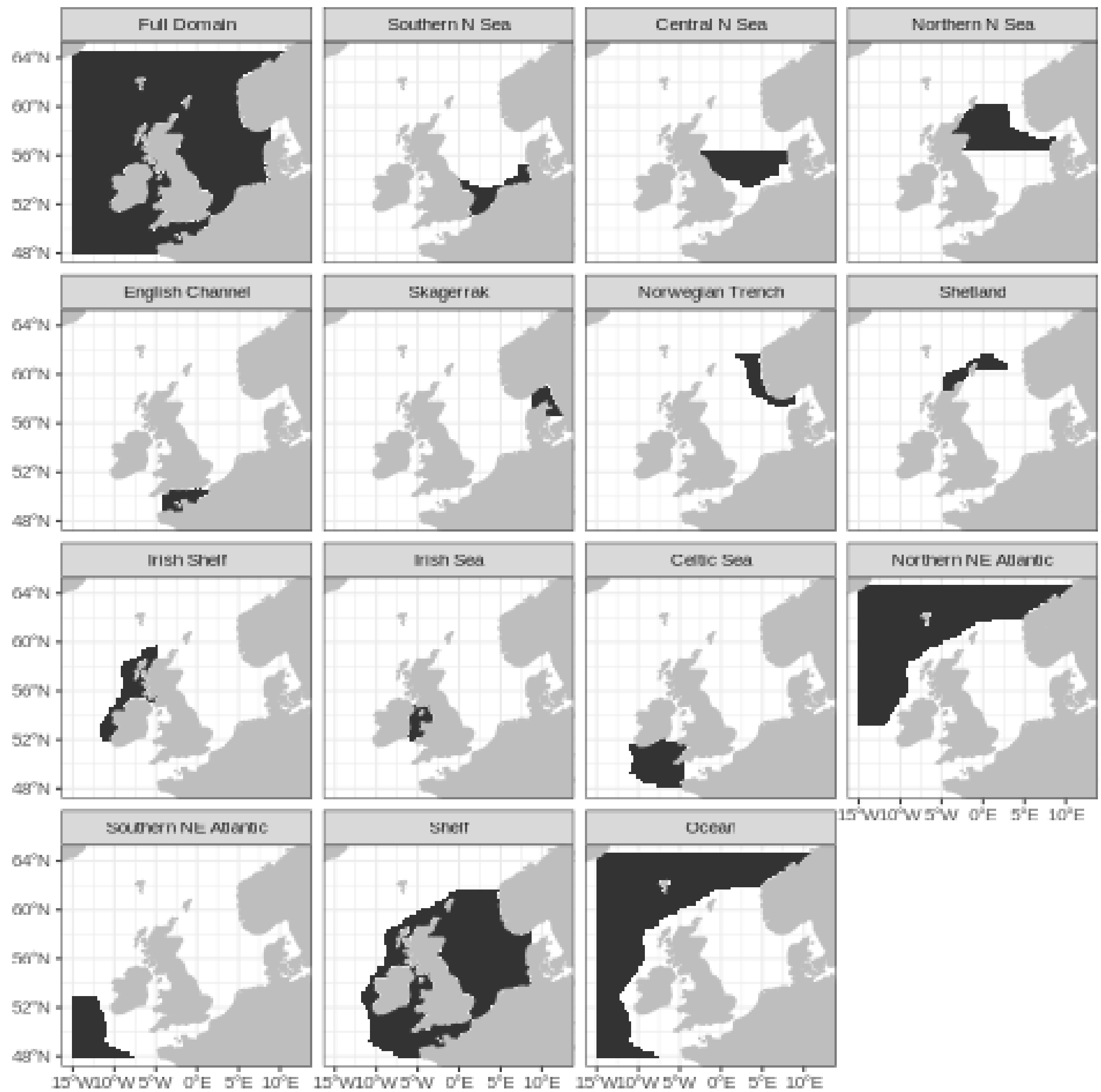


Figure 19.5: Regions used for validation of sea surface salinity.

Time series were constructed comparing the monthly mean of the spatial average sea surface salinity in each region. The spatial average was calculated using the mean of all grid cells within each region, accounting for grid cell area.

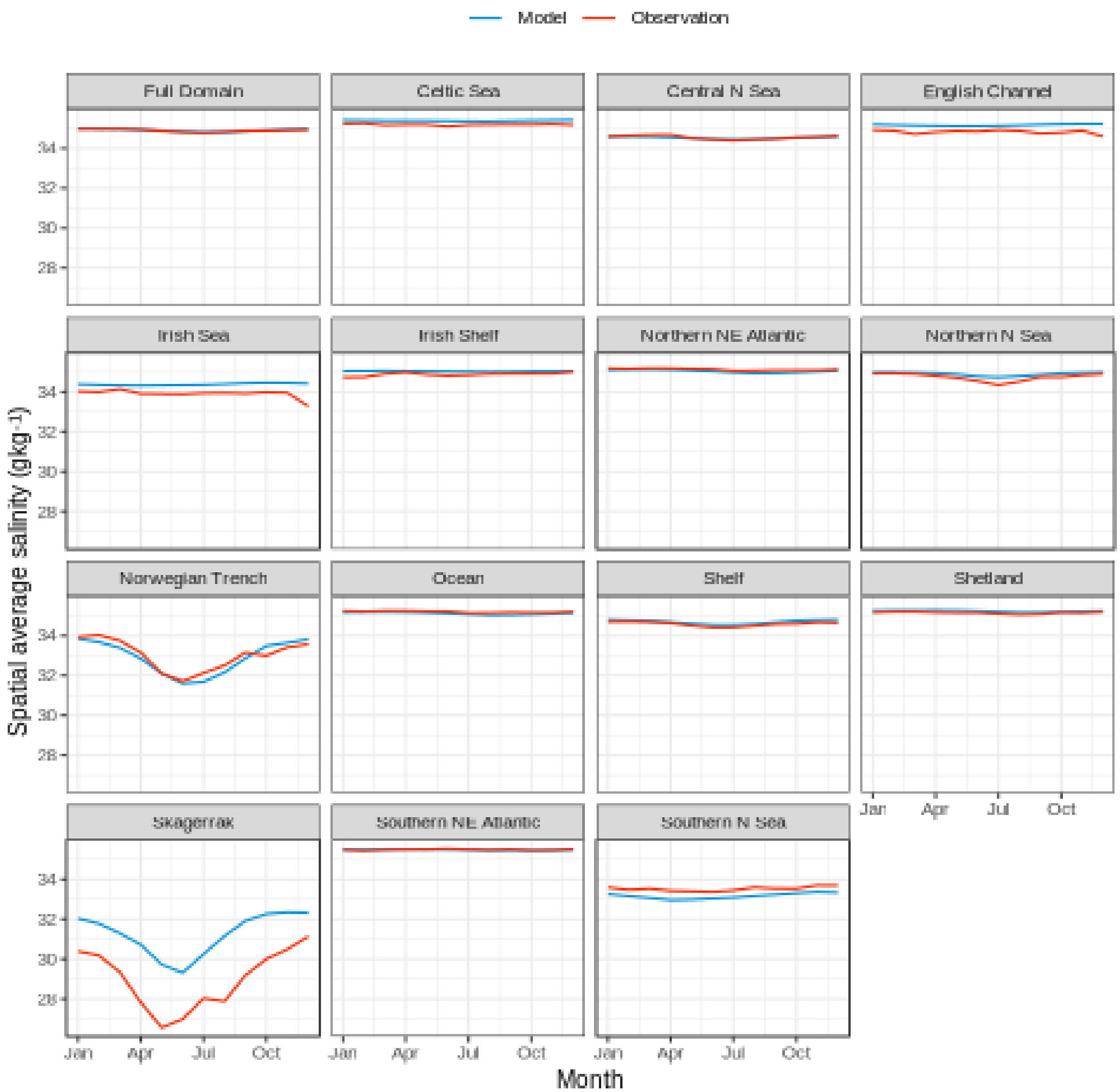


Figure 19.6: Seasonal cycle of sea surface salinity for model and observations for each region. The spatial average is taken over the region.

The table below shows the average bias of sea surface salinity in each region. The bias is calculated as the modelled value minus the observed value. A positive bias indicates that the model overestimates the observed value, while a negative bias indicates that the model underestimates the observed value.

Region	Temporal correlation	Bias	RMSD
Full Domain	0.44	0.01	0.47
Celtic Sea	0.24	0.20	0.43
Central North Sea	0.60	-0.03	0.18
Channel	0.14	0.34	0.54
Irish Sea	-0.05	0.47	0.73
Irish Shelf	0.12	0.15	0.27
Northern North East Atlantic	0.54	-0.09	0.23
Northern North Sea	0.69	0.15	0.35
Norwegian Trench	0.88	-0.11	0.69
Ocean	0.47	-0.08	0.22
Shelf	0.43	0.10	0.64
Shetland	0.50	0.08	0.13
Skagerrak	0.79	2.26	4.16
Southern North East Atlantic	0.13	-0.01	0.08
Southern North Sea	0.35	-0.36	1.79

Table 19.1: Summary of performance of the model sea surface salinity in each region. The bias (g kg^{-1}) column represents the spatial average of the annual mean modelled value minus the observed value. The temporal correlation column represents the spatial mean of the temporal correlation between the model and observations per grid cell.

19.5 Can the model reproduce spatial patterns of sea surface salinity?

The ability of the model to reproduce spatial patterns of sea surface salinity was assessed by comparing the modelled and observed salinity at each grid cell. We calculated the Pearson correlation coefficient between the modelled and observed salinity at each grid cell.

This was carried out monthly and using the annual mean in each grid cell

Time period	r
Annual mean	0.78
Jan	0.72
Feb	0.74
Mar	0.74
Apr	0.75
May	0.77
Jun	0.84
Jul	0.82
Aug	0.77
Sep	0.76
Oct	0.78
Nov	0.76
Dec	0.70

Table 19.2: Pearson correlation coefficient between modelled and observed sea surface salinity at each grid cell. The correlation was calculated monthly and using the annual mean in each grid cell.

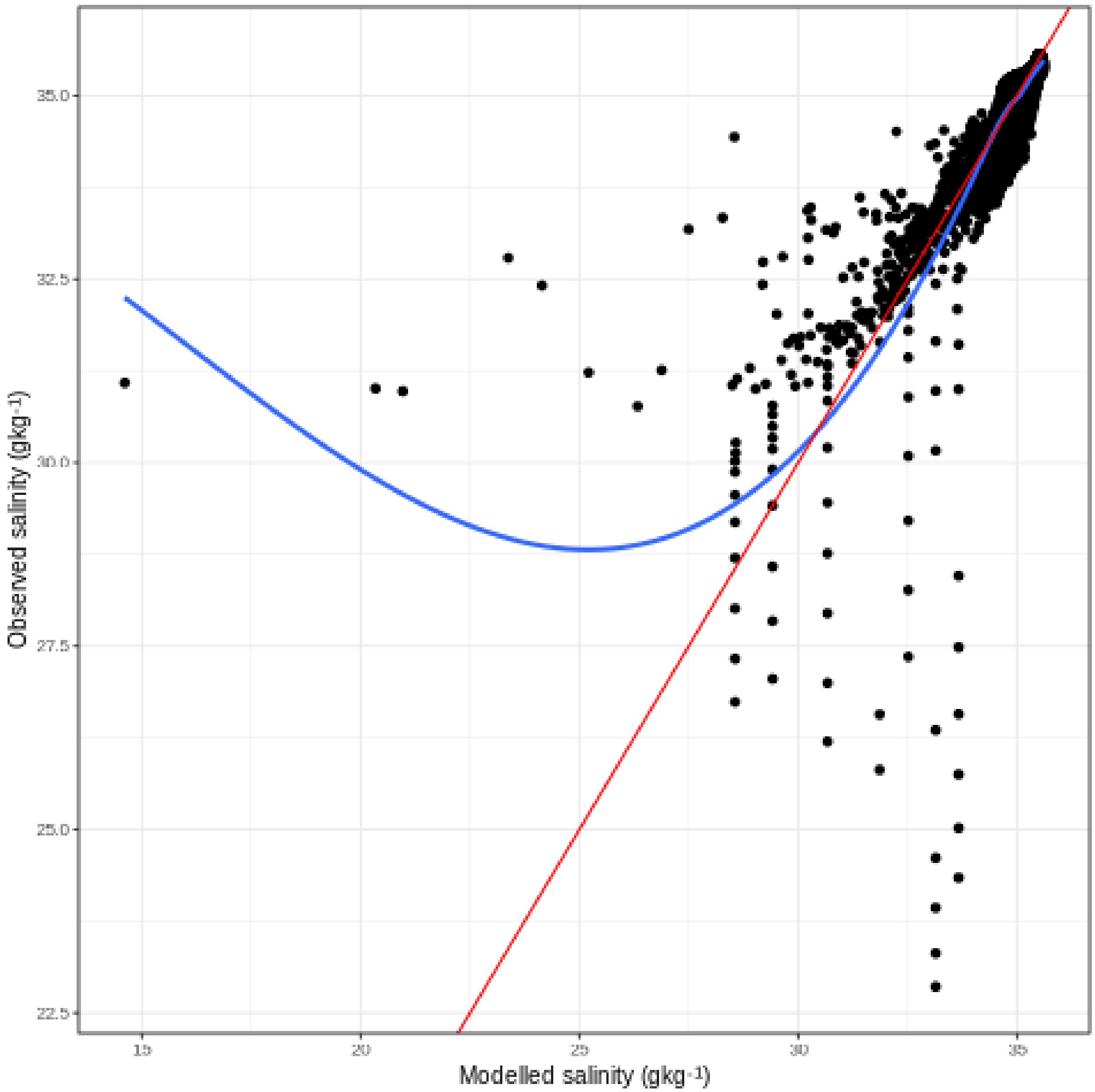


Figure 19.7: Scatter plot of modelled and observed annual average sea surface salinity in grid cells. The red line is the 1:1 line. The blue line is the linear regression of the modelled and observed values. The slope of the blue line is the slope of the regression line.

19.6 Data Sources for validation of salinity

Hinrichs, Iris; Gouretski, Viktor; Paetsch, Johannes; Emeis, Kay; Stammer, Detlef (2017). North Sea Biogeochemical Climatology (Version 1.1).

URL: https://www.cen.uni-hamburg.de/en/icdc/data/ocean/nsbc.html

SEA SURFACE SILICATE VALIDATION USING GRIDDED OBSERVATIONS FROM NSBC

We used version 1.1 of the **North Sea Biogeochemical Climatology** (NSBC) to validate **sea surface silicate**. NSBC is a monthly climatology that covers the region 47°-65°N and 15°W-15°E. The data is made up of observations over the period 1960-2014. For validation purposes we only used the level 3 data, which a gridded monthly climatology at a spatial resolution of 1/4°. The data can be download from NSBC.

Matchup procedure: The model and observations were matched up as follows. First, the model dataset was cropped by a small amount to make sure cells close to the boundary were removed. The model was then regridded to the observational grid if the observational grid was coarser using nearest neighbour. Only grid cells with model and observational data were maintained. The following model output was used to compare with the observational values: **N5_s**.

20.1 Baseline climatologies of sea surface silicate

Climatologies of model and observational sea surface silicate concentration are shown in the figures below. The model climatology is calculated using the years **1986-2017**.

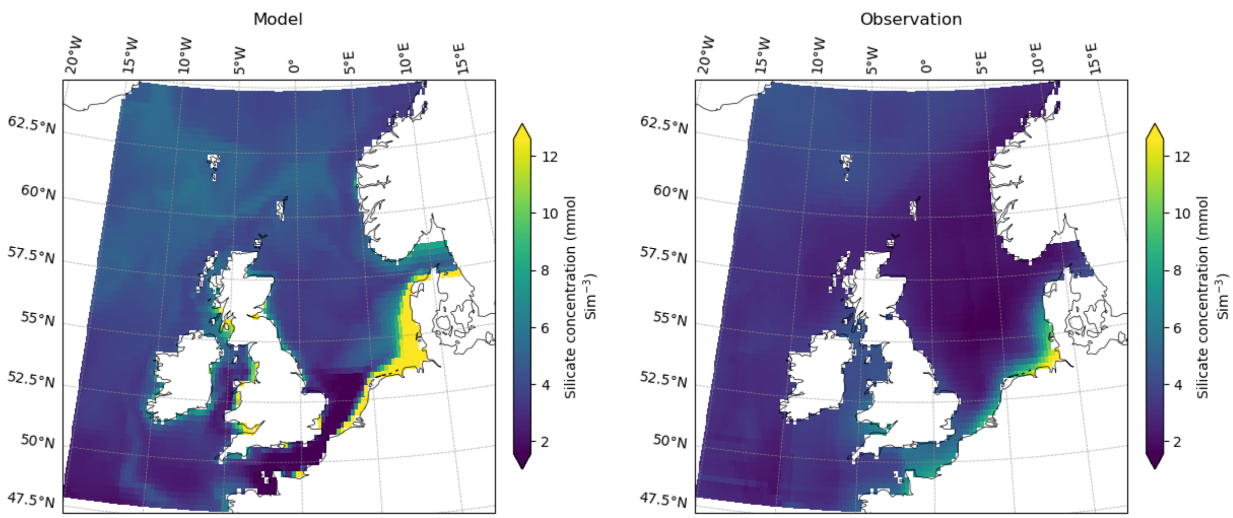


Figure 20.1: Annual average surface silicate concentration from the model (1986-2017) and observations. Data is limited to the 2nd and 98th percentile of the combined model and observational data. Arrows indicate that values can exceed the colourbar limits.

20.2 Assessing model bias for surface silicate concentration

Figure 20.2 shows the average bias of surface silicate concentration simulated by the model. A positive bias indicates that the model overestimates the observation, while a negative bias indicates that the model overpredicts the observation.

The spatial average bias of surface silicate concentration is 1.32 mmol Sim^{-3}. Overall, the model overestimates the observations in 83.1% of the model domain.

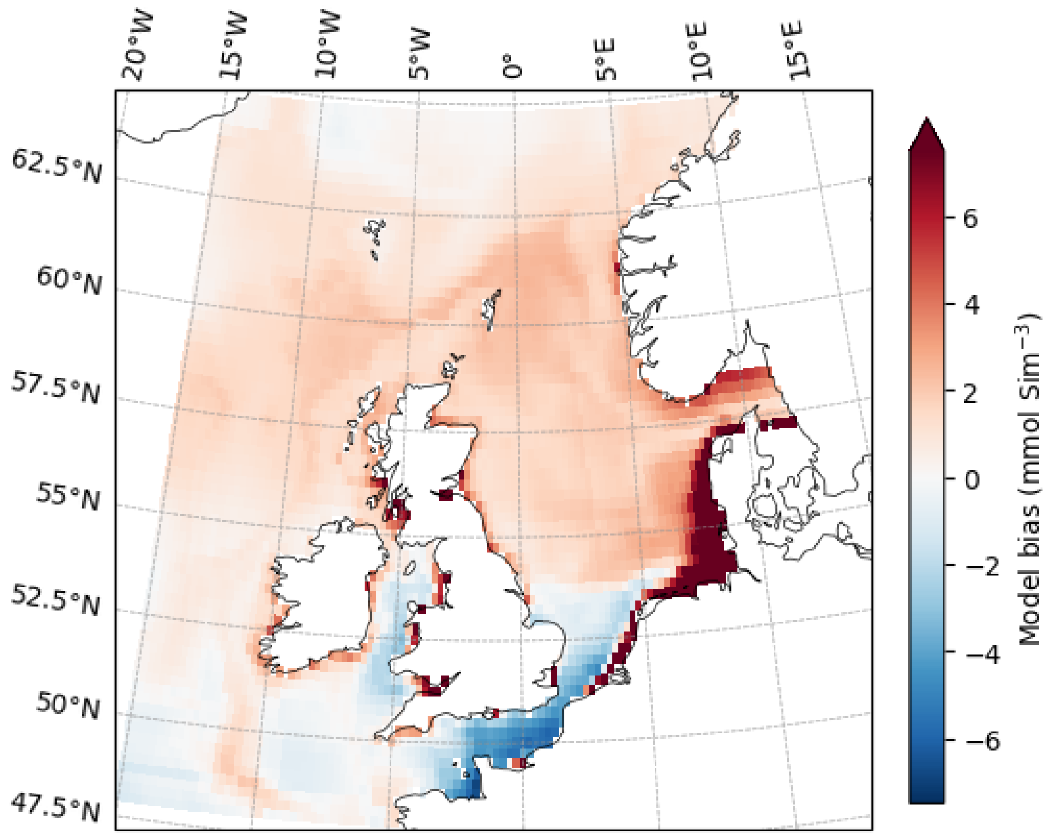


Figure 20.2: Bias of surface silicate concentration from the model. A positive bias indicates that the model overestimates the observation. For clarity, the colourbar is limited to the 2nd and 98th percentile of the data.

20.3 Can the model reproduce seasonality of sea surface silicate concentration?

The ability of the model to reproduce seasonality of sea surface silicate concentration was assessed by comparing the modelled and observed seasonal cycle of silicate concentration. First, we derive a monthly climatology for the model data. Then, we calculate the Pearson correlation coefficient between the modelled and observed silicate concentration at each grid cell.

Note: we are only assessing the ability of the model to reproduce the ability of the model to reproduce seasonal changes, not long-term trends.

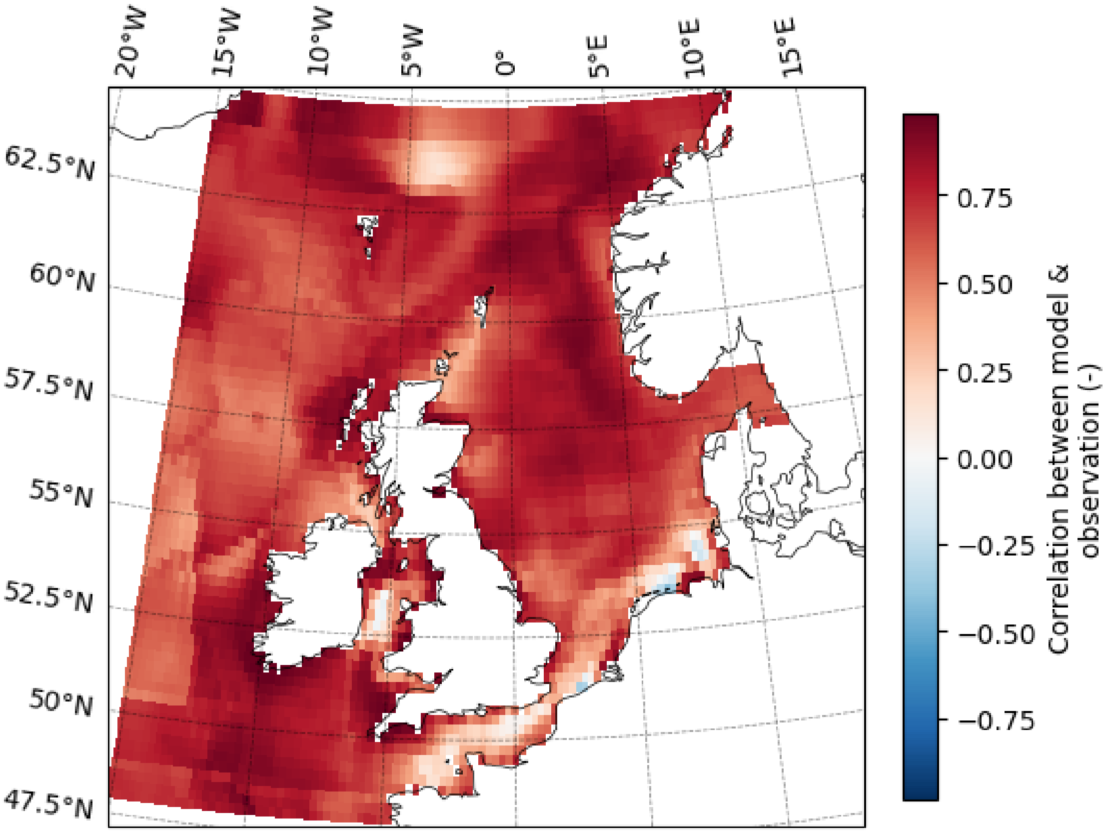
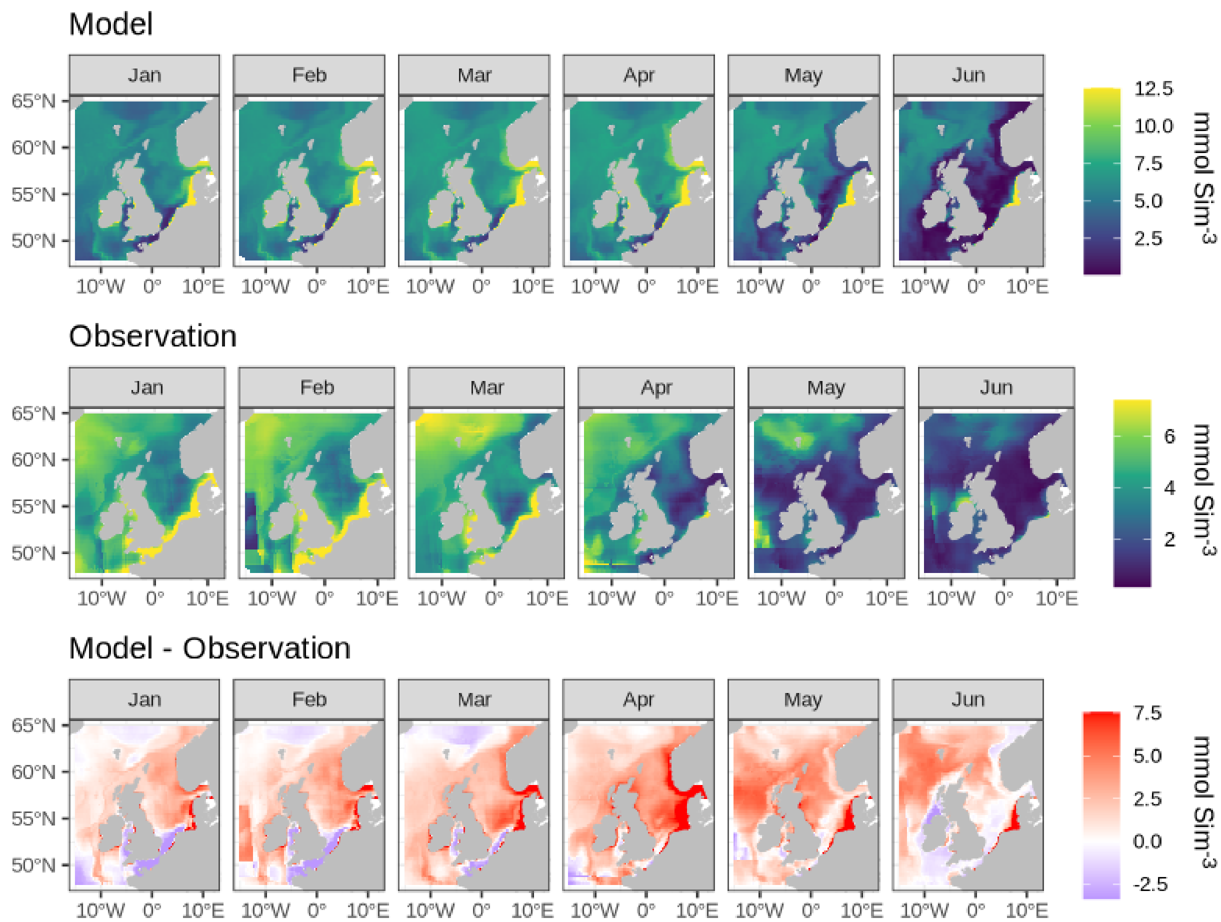


Figure 20.3: Seasonal temporal correlation between model and observations for surface silicate concentration. This is the Pearson correlation coefficient between climatology monthly mean values in the model and observations.

The seasonal cycles of simulated and observed silicate concentration are compared in Figure 20.4 below. This figure shows the model and observation average in each month of the year, and the differences between the two each month


Model
Jan
Feb
Mar
Apr
May
Jun
65°N
60°N
55°N
50°N
10°W
0°
10°E
12.5
10.0
7.5
5.0
2.5
mmol Sim-3
Observation
6
4
2
Model - Observation
7.5
5.0
2.5
0.0
-2.5

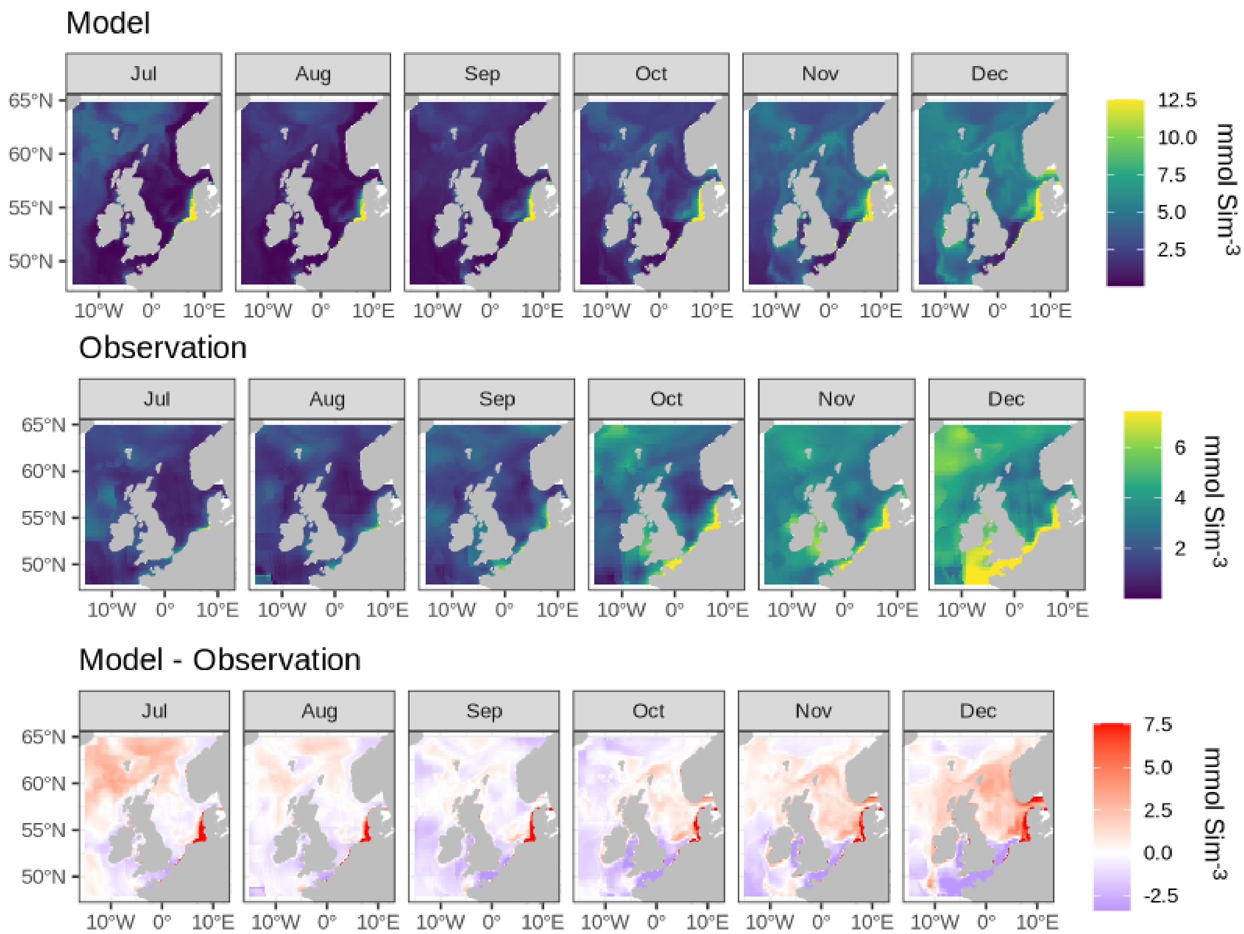


Figure 20.4: Monthly mean surface silicate concentration for the model, observation and the difference between model and observations. For clarity, the maximum values are capped to the 98th percentiles.

20.4 Regional assessment of model performance for sea surface silicate concentration

We assessed the regional performance of the model by comparing the model with observations in a number of regions. The regions considered are mapped below.

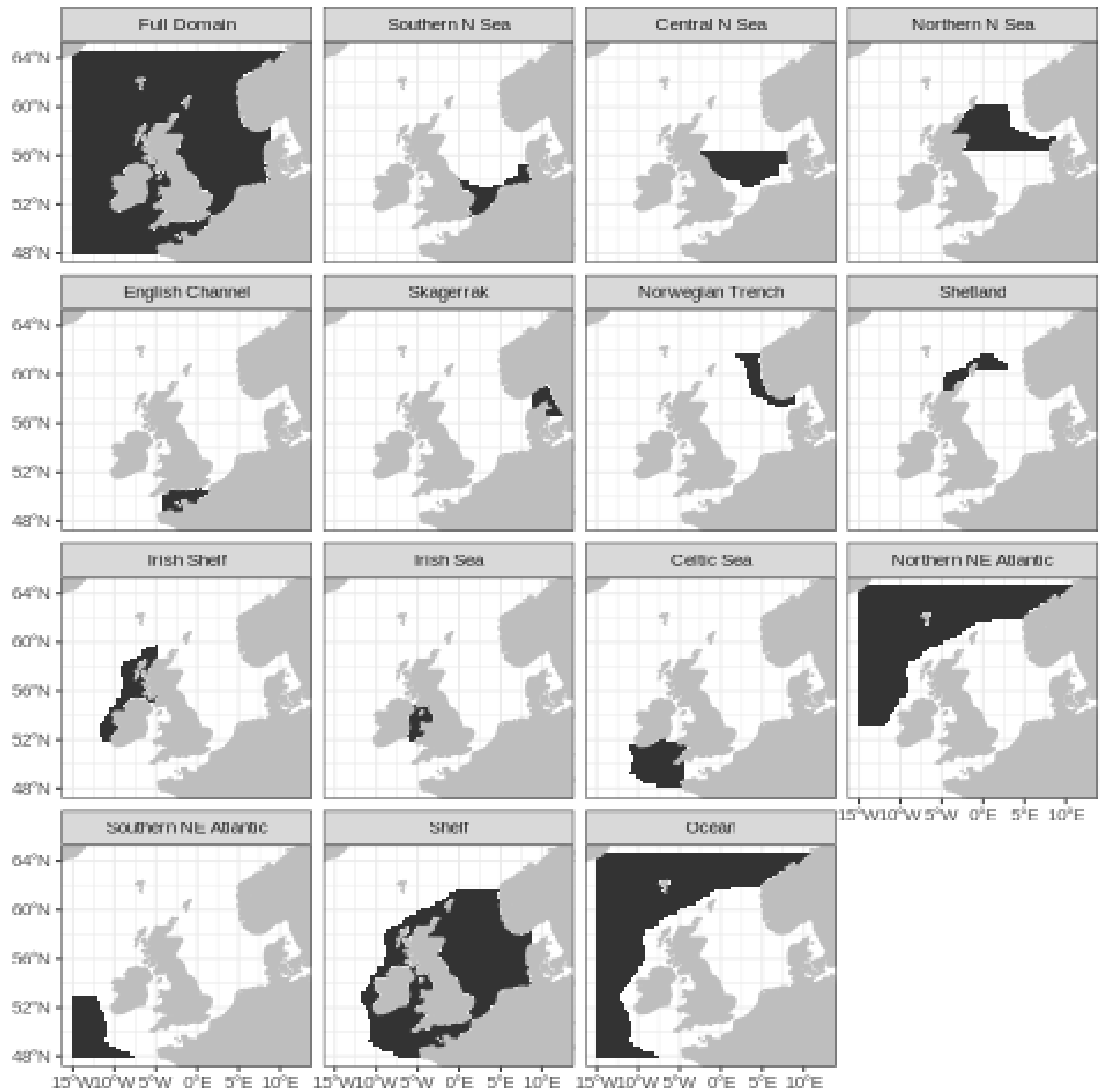


Figure 20.5: Regions used for validation of sea surface silicate concentration.

Time series were constructed comparing the monthly mean of the spatial average sea surface silicate concentration in each region. The spatial average was calculated using the mean of all grid cells within each region, accounting for grid cell area.

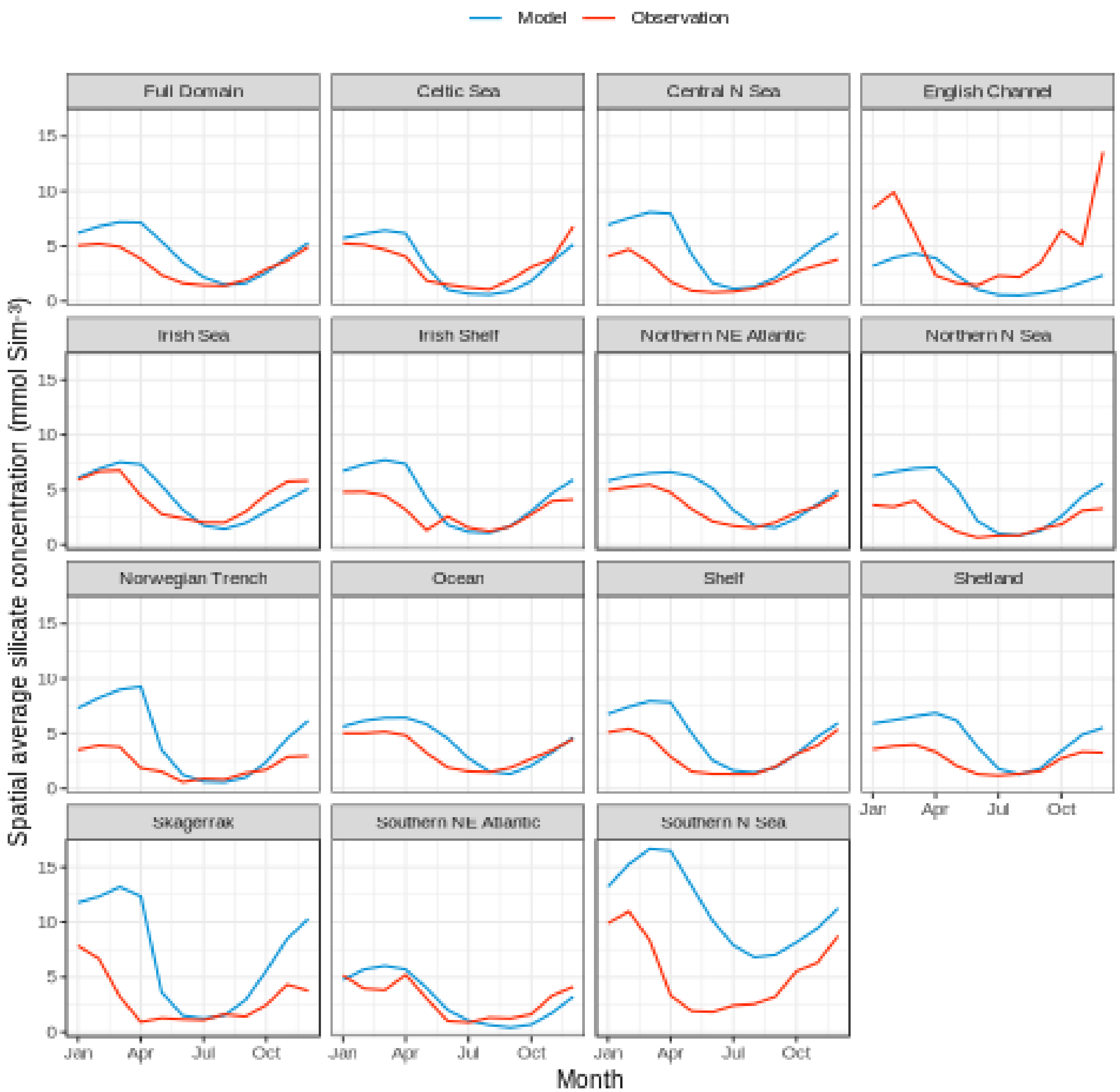


Figure 20.6: Seasonal cycle of sea surface silicate concentration for model and observations for each region. The spatial average is taken over the region.

The table below shows the average bias of sea surface silicate concentration in each region. The bias is calculated as the modelled value minus the observed value. A positive bias indicates that the model overestimates the observed value, while

a negative bias indicates that the model underestimates the observed value.

Region	Temporal correlation	Bias	RMSD
Full Domain	0.69	1.18	4.00
Celtic Sea	0.80	0.07	1.96
Central North Sea	0.70	2.22	4.14
Channel	0.38	-3.12	5.19
Irish Sea	0.59	0.13	2.92
Irish Shelf	0.75	1.36	2.45
Northern North East Atlantic	0.69	0.99	1.83
Northern North Sea	0.75	1.94	2.94
Norwegian Trench	0.79	2.34	3.51
Ocean	0.70	0.83	1.81
Shelf	0.69	1.52	5.51
Shetland	0.77	1.89	2.39
Skagerrak	0.63	4.09	6.11
Southern North East Atlantic	0.73	0.10	1.68
Southern North Sea	0.43	5.88	16.07

Table 20.1: Summary of performance of the model sea surface silicate concentration in each region. The bias (mmol Sim-3) column represents the spatial average of the annual mean modelled value minus the observed value. The temporal correlation column represents the spatial mean of the temporal correlation between the model and observations per grid cell.

20.5 Can the model reproduce spatial patterns of sea surface silicate concentration?

The ability of the model to reproduce spatial patterns of sea surface silicate concentration was assessed by comparing the modelled and observed silicate concentration at each grid cell. We calculated the Pearson correlation coefficient between the modelled and observed silicate concentration at each grid cell.

This was carried out monthly and using the annual mean in each grid cell

Time period	*r*
Annual mean	0.61
Jan	0.57
Feb	0.57
Mar	0.56
Apr	0.09
May	0.21
Jun	0.45
Jul	0.61
Aug	0.58
Sep	0.53
Oct	0.55
Nov	0.68
Dec	0.25

Table 20.2: Pearson correlation coefficient between modelled and observed sea surface silicate concentration at each grid cell. The correlation was calculated monthly and using the annual mean in each grid cell.

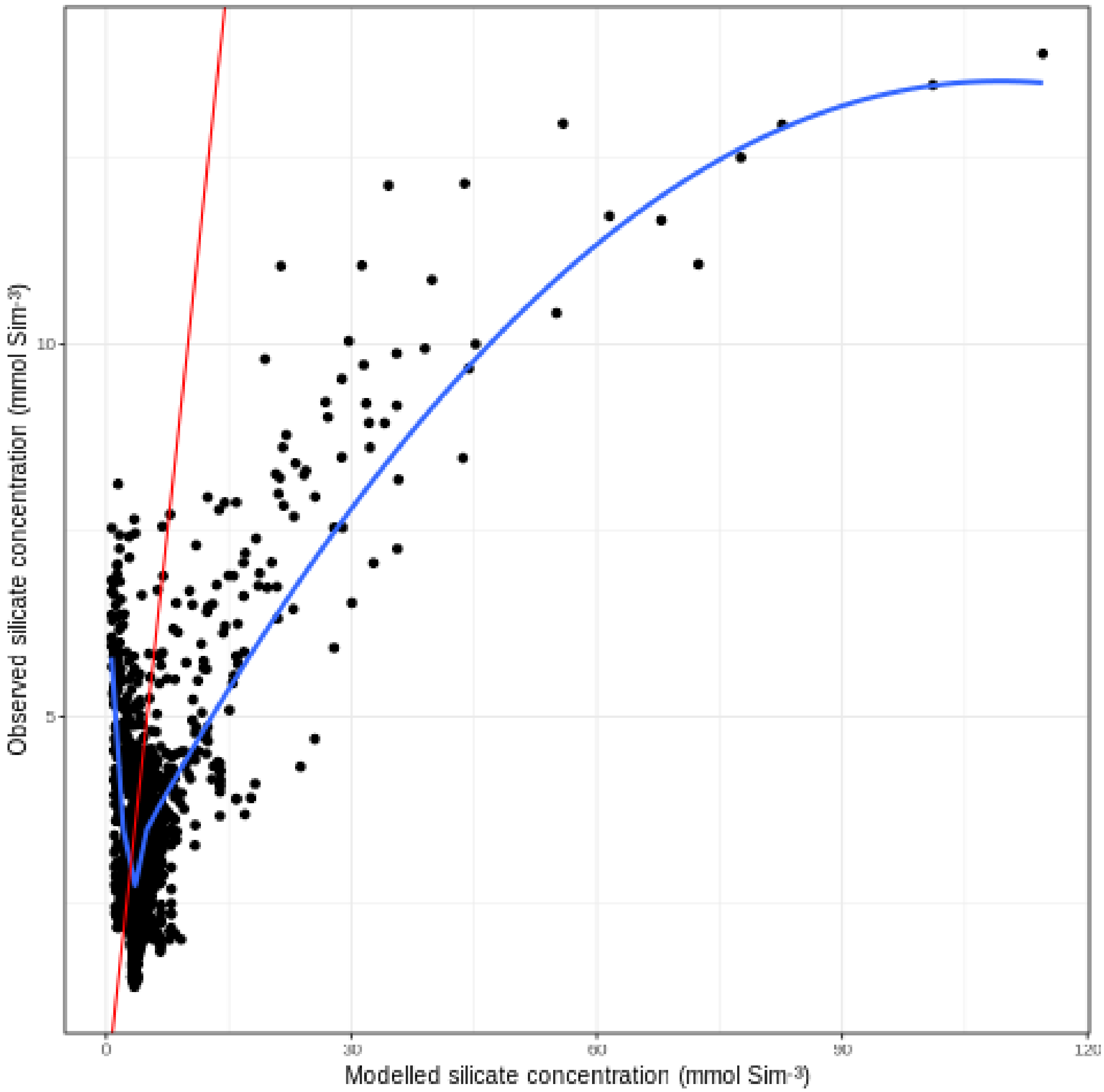


Figure 20.7: Scatter plot of modelled and observed annual average sea surface silicate concentration in grid cells. The red line is the 1:1 line. The blue line is the linear regression of the modelled and observed values. The slope of the blue line is the slope of the regression line.

20.6 Data Sources for validation of silicate concentration

Hinrichs, Iris; Gouretski, Viktor; Paetsch, Johannes; Emeis, Kay; Stammer, Detlef (2017). North Sea Biogeochemical Climatology (Version 1.1).

URL: https://www.cen.uni-hamburg.de/en/icdc/data/ocean/nsbc.html

CHAPTER

TWENTYONE

VALIDATION OF BOTTOM TEMPERATURE USING POINT OBSERVATIONS

21.1 Performance of model near-bottom temperature

Near-bottom values of temperature were extracted from International Council for the Exploration of the Sea (ICES) bottle and CTD data. The near-bottom was defined as observations **within 2m of the seabed**. This was interpolated to the observational grid using the $GEBCO_2$ bathymetry dataset. Model values were interpolated to the observational dataset's longitudes and latitudes using 3D interpolation. **Note:** this analysis has been restricted to observations on the shelf region. The model output was matched up with the observational data with model output from the years **1986 to 2017**. The model output was matched up precisely with the observational data for each day of the year in the years with data in both model and observations.

In total there were 8,733 values extracted from the observational database. The map below shows the locations of the matched up data for temperature.

The following model output was used to compare with observational values: **Y3_c/(Y2_c + Y3_c)**.

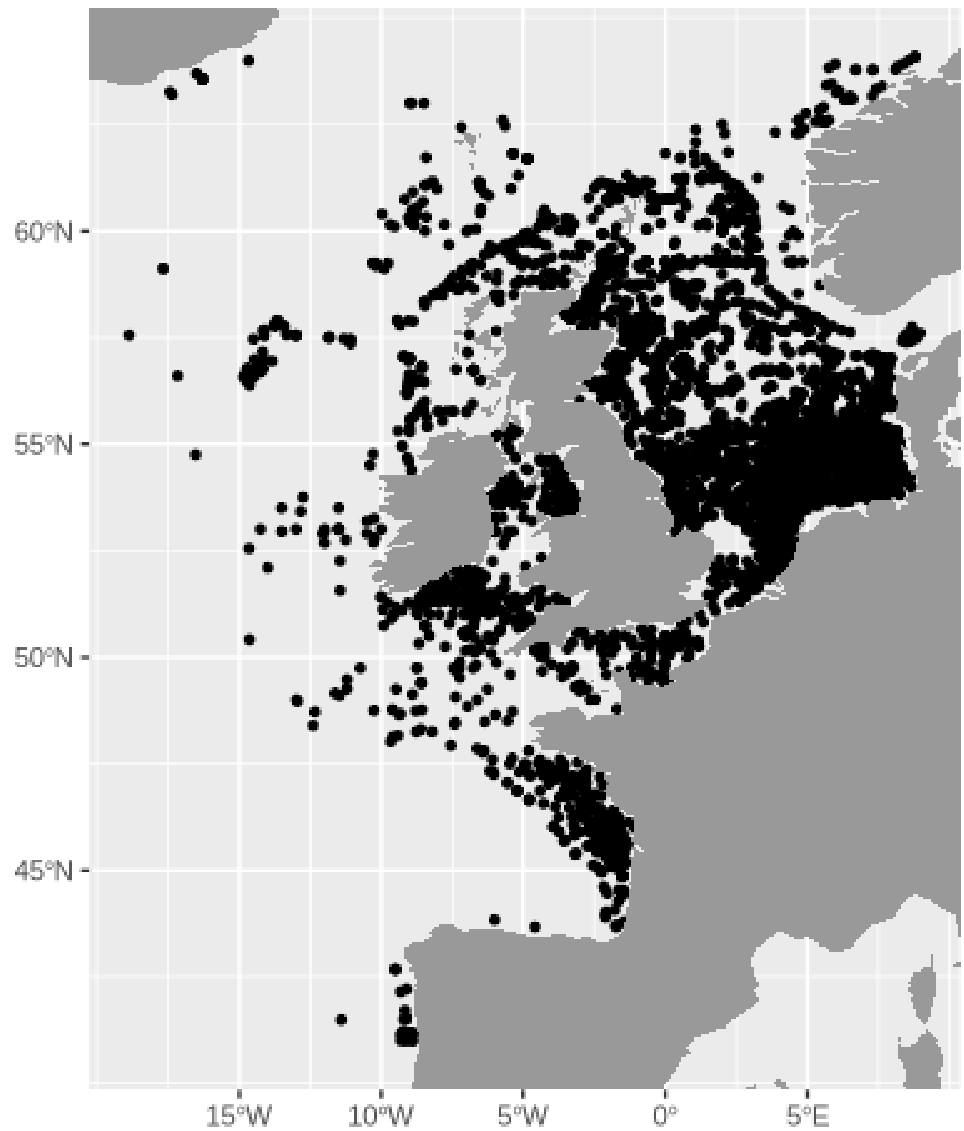


Figure 21.1: Locations of matchups between simulated and observed temperature near the bottom of the water column.

The number of observations in each month ranged from 447 in October to 1,345 in August. Figure 21.2 below shows the distribution of observations in each month.

Figure 21.2 below shows the bias between the model and observational data for temperature. The bias is calculated as the model value minus the observational value, and it is shown for each month of the year.

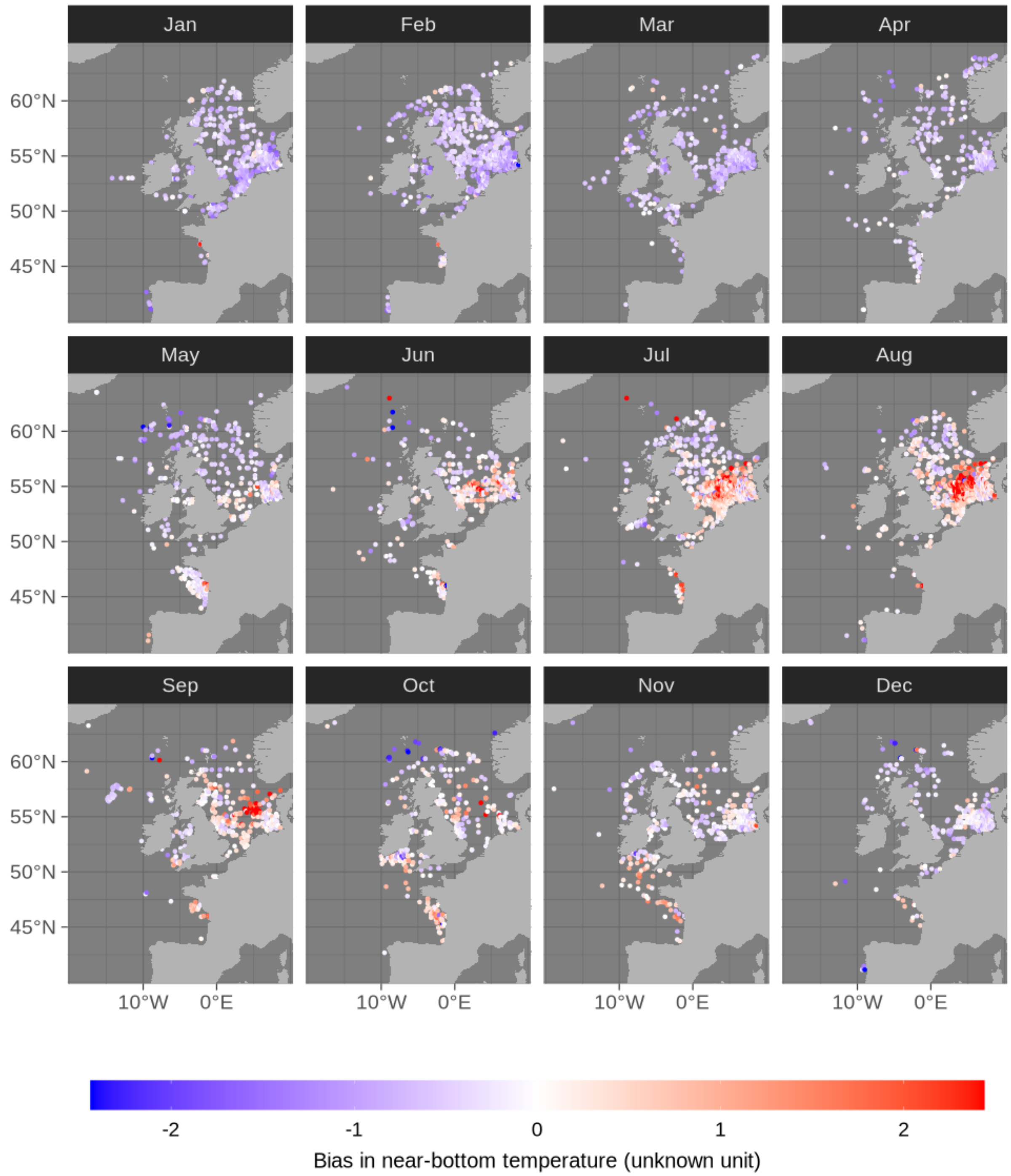


Figure 21.2: Bias in near-bottom temperature. The bias is calculated as model-observation. The colour scale is from blue (negative bias) to red (positive bias). The colour scale is capped at the 98th percentile of the absolute bias. This is to avoid a few extreme outliers from dominating the colour scale. **Note:** values have been binned and averaged to the resolution of the model.

Figures 21.3 and 21.4 show the distribution of near-bottom temperature observations in the model and observational datasets. This is shown for each month of the year (Figure 21.3) and for the entire year (Figure 21.4).

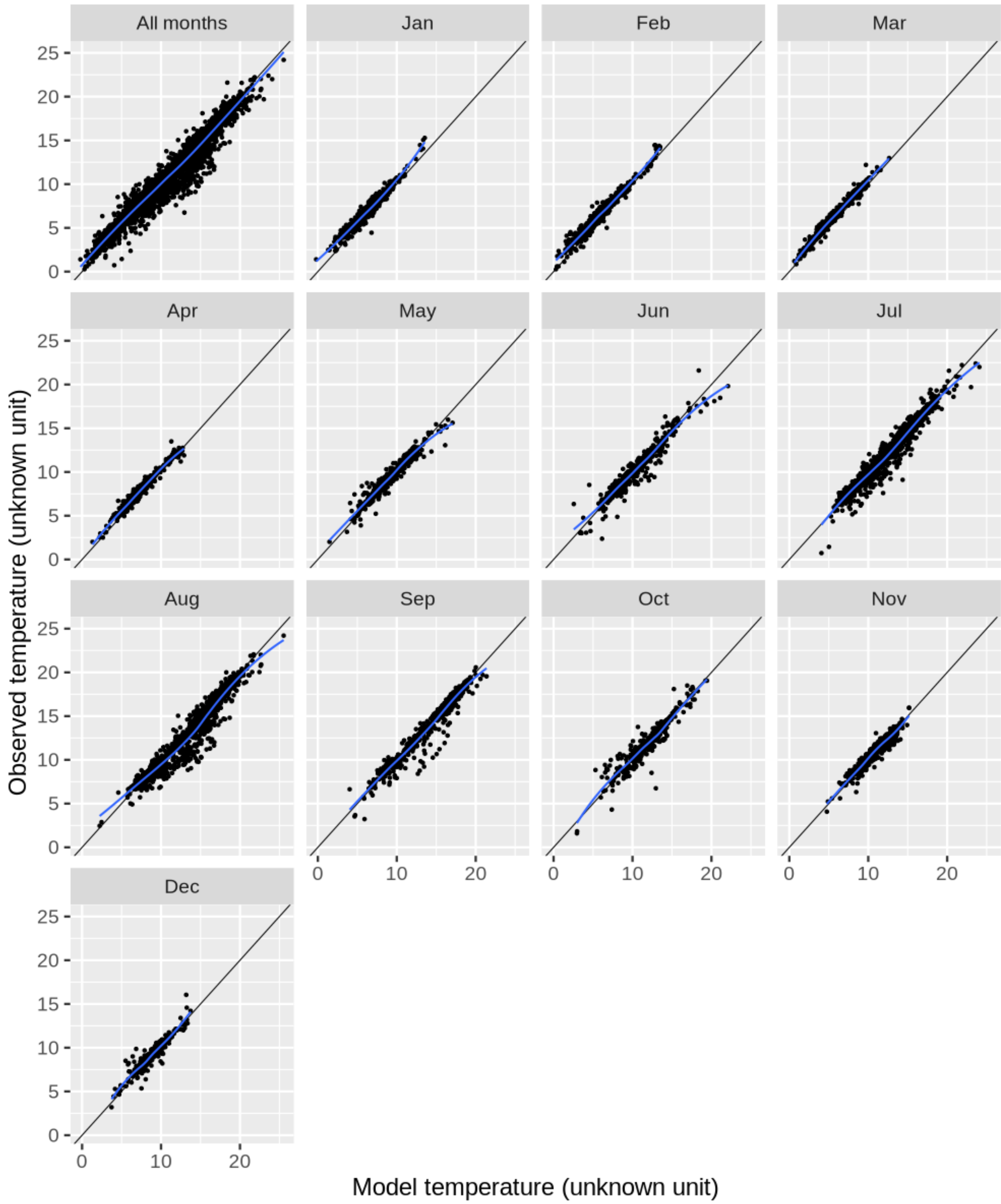


Figure 21.3: Simulated versus observed temperature near the bottom of the water column. The blue curve is a generalized additive model (GAM) fit to the data, and the black line represents 1-1 relationship between the simulation and observations. The data has been averaged per model grid cell.

21.2 Summary statistics for near-bottom temperature

The overall ability of the model to predict the observed temperature was assessed by calculating the average bias, the root mean square deviation (RMSD) and the correlation coefficient (R). The bias was calculated as the model value minus the observed value. The RMSD was calculated as the square root of the mean squared deviation. The correlation coefficient was calculated as the Pearson correlation coefficient between the model and observed values.

This was calculated for each month and for the entire dataset. The results are shown in the tables below.

Time period	Bias	RMSD	*r*	Number of observations
All	-0.09	0.83	0.98	8,733
Jan	-0.60	0.75	0.97	769
Feb	-0.60	0.73	0.98	976
Mar	-0.61	0.73	0.98	546
Apr	-0.43	0.61	0.98	529
May	-0.21	0.62	0.97	782
Jun	0.13	0.89	0.97	557
Jul	0.29	0.93	0.97	1,130
Aug	0.45	1.13	0.97	1,345
Sep	0.23	0.88	0.97	555
Oct	0.02	0.83	0.95	447
Nov	-0.10	0.50	0.95	639
Dec	-0.29	0.63	0.93	458

Table 21.1: Average bias (unknown unit) and root-mean square deviation (unknown unit) for the model's near-bottom temperature for each month. The bias is calculated as model-observation. The average bias is calculated as the mean of the monthly biases.

A linear regression analysis of modelled and observed temperature was performed. The modelled temperature was used as the independent variable and the observed temperature was used as the dependent variable. The results are shown in the table below.

The regression was carried out using the Python package statsmodels.

Period	Slope	Intercept	R^2	p-value
All	0.91	1.03	0.96	0.00
Jan	0.95	0.93	0.95	0.00
Feb	0.94	0.93	0.95	0.00
Mar	0.97	0.77	0.97	0.00
Apr	0.95	0.85	0.96	0.00
May	0.92	1.00	0.95	0.00
Jun	0.92	0.71	0.93	0.00
Jul	0.95	0.30	0.94	0.00
Aug	0.96	0.03	0.93	0.00
Sep	0.97	0.26	0.95	0.00
Oct	0.92	0.97	0.88	0.00
Nov	0.93	0.90	0.91	0.00
Dec	0.92	1.08	0.89	0.00

Table 21.2: Linear regression analysis of modelled and observed temperature. The modelled temperature was used as the independent variable and the observed temperature was used as the dependent variable. The slope and intercept of the regression line are shown, along with the R^2 value and the p-value of the slope. Note: only months with sufficient values for a regression are shown.

21.3 Data Sources for validation of temperature

ICES Data Portal, Dataset on Ocean HydroChemistry, Extracted March 3, 2023. ICES, Copenhagen

CHAPTER

TWENTYTWO

SEA SURFACE TEMPERATURE VALIDATION USING GRIDDED OBSERVATIONS FROM OSTIA

Temperature was validated using the OSTIA sea surface temperature dataset. The validation was performed by comparing the modelled temperature with the OSTIA data for the same time and location. The OSTIA data was downloaded from the Copernicus Marine Environment Monitoring Service (CMEMS) catalogue. A description of the dataset is available here.

Matchup procedure: The model and observations were matched up as follows. First, the model dataset was cropped by a small amount to make sure cells close to the boundary were removed. The model was then regridded to the observational grid if the observational grid was coarser using nearest neighbour. Only grid cells with model and observational data were maintained. The following model output was used to compare with the observational values: **votemper**.

22.1 Baseline climatologies of sea surface temperature

Climatologies of model and observational sea surface temperature are shown in the figures below. The model climatology is calculated using the years **1986-2017**.

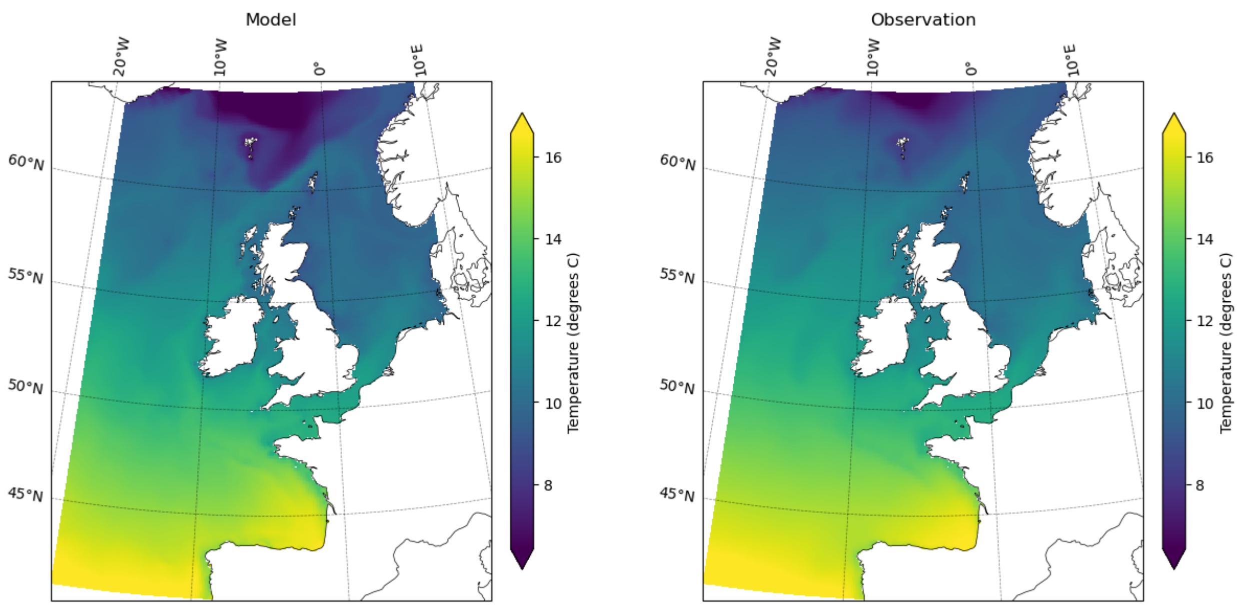


Figure 22.1: Annual average surface temperature from the model (1986-2017) and observations. Data is limited to the 2nd and 98th percentile of the combined model and observational data. Arrows indicate that values can exceed the colourbar limits.

22.2 Assessing model bias for surface temperature

Figure 22.2 shows the average bias of surface temperature simulated by the model. A positive bias indicates that the model overestimates the observation, while a negative bias indicates that the model overpredicts the observation.

The spatial average bias of surface temperature is -0.30 degrees C. Overall, the model underestimates the observations in 88.1% of the model domain.

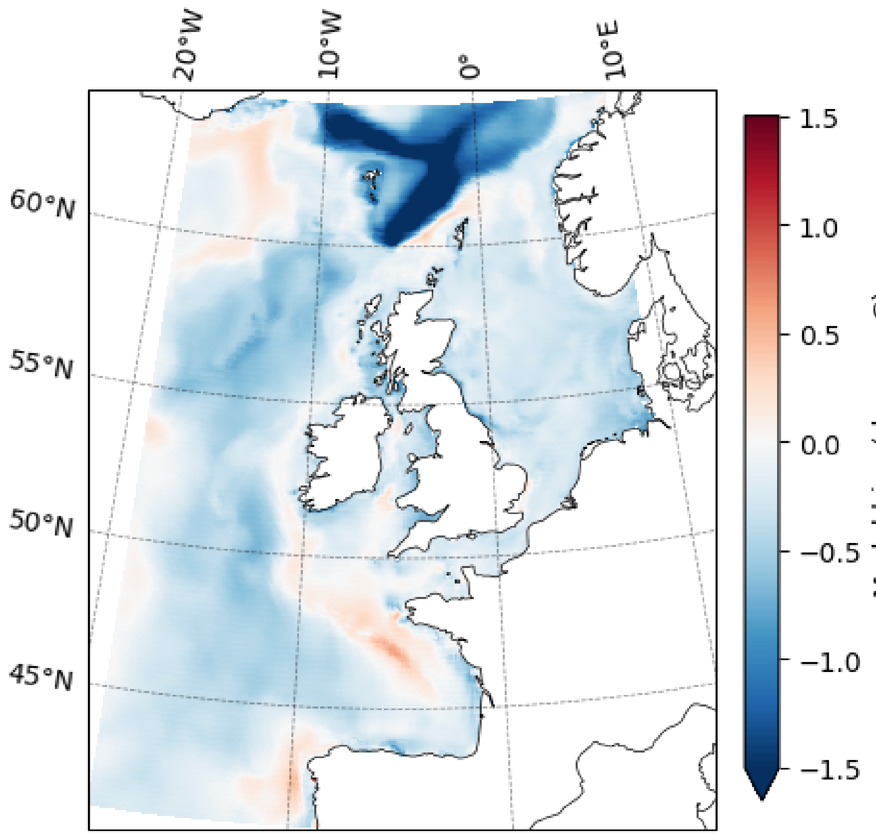


Figure 22.2: Bias of surface temperature from the model. A positive bias indicates that the model overestimates the observation. For clarity, the colourbar is limited to the 2nd and 98th percentile of the data.

22.3 Can the model reproduce seasonality of sea surface temperature?

The ability of the model to reproduce seasonality of sea surface temperature was assessed by comparing the modelled and observed seasonal cycle of temperature. First, we derive a monthly climatology for the model data. Then, we calculate the Pearson correlation coefficient between the modelled and observed temperature at each grid cell.

Note: we are only assessing the ability of the model to reproduce the ability of the model to reproduce seasonal changes, not long-term trends.

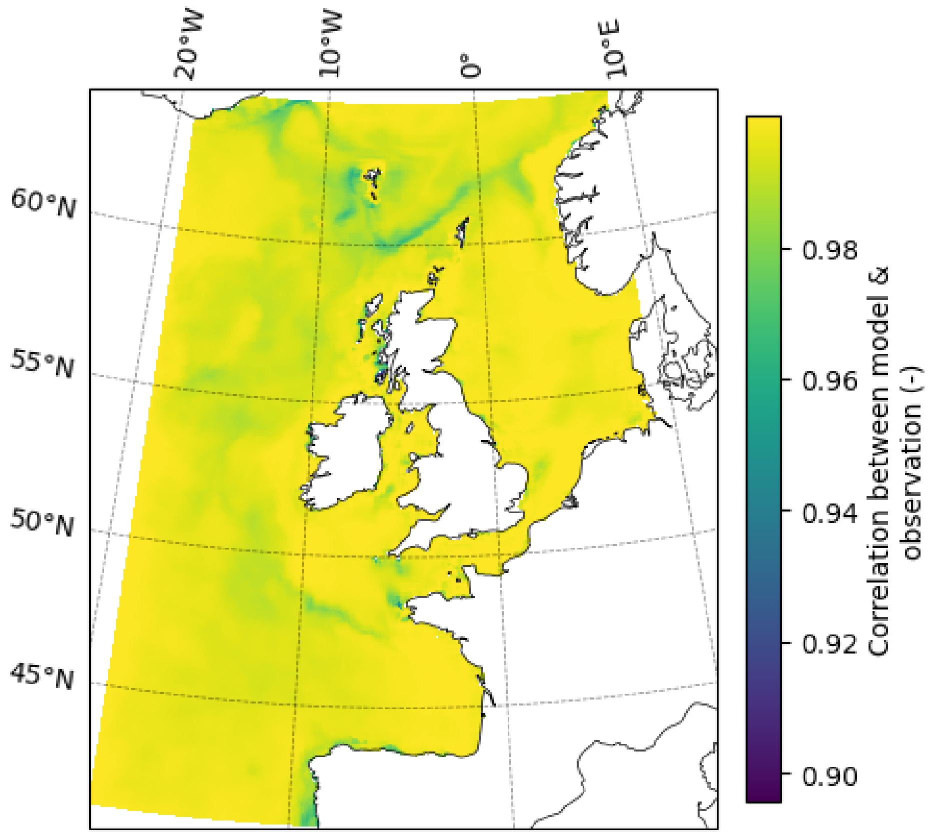
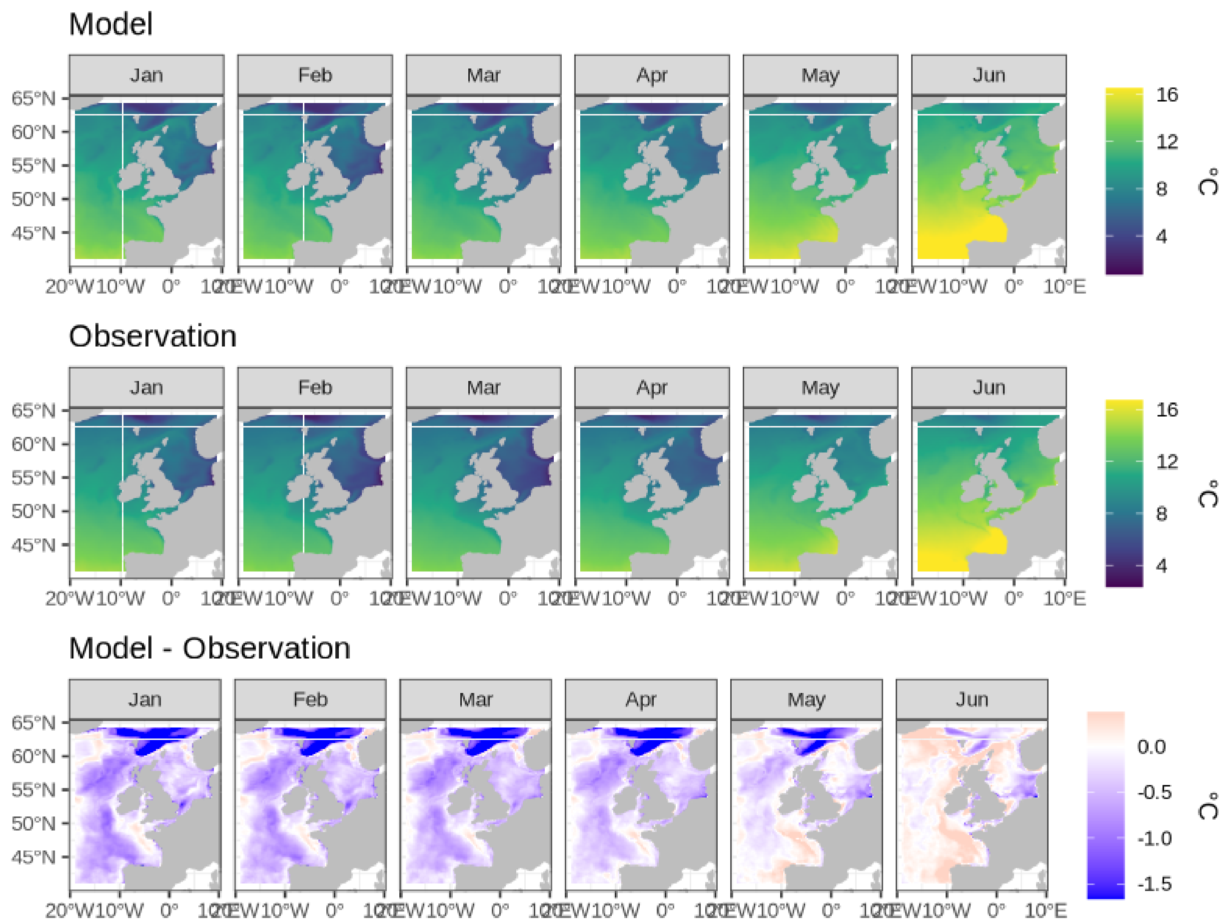


Figure 22.3: Seasonal temporal correlation between model and observations for surface temperature. This is the Pearson correlation coefficient between climatology monthly mean values in the model and observations.

The seasonal cycles of simulated and observed temperature are compared in Figure 22.4 below. This figure shows the model and observation average in each month of the year, and the differences between the two each month


Model
Jan
Feb
Mar
Apr
May
Jun
65°N
60°N
55°N
50°N
45°N
20°W
10°W
0°
10°E
16
12
8
4
°C
Observation
Model - Observation
0.0
-0.5
-1.0
-1.5

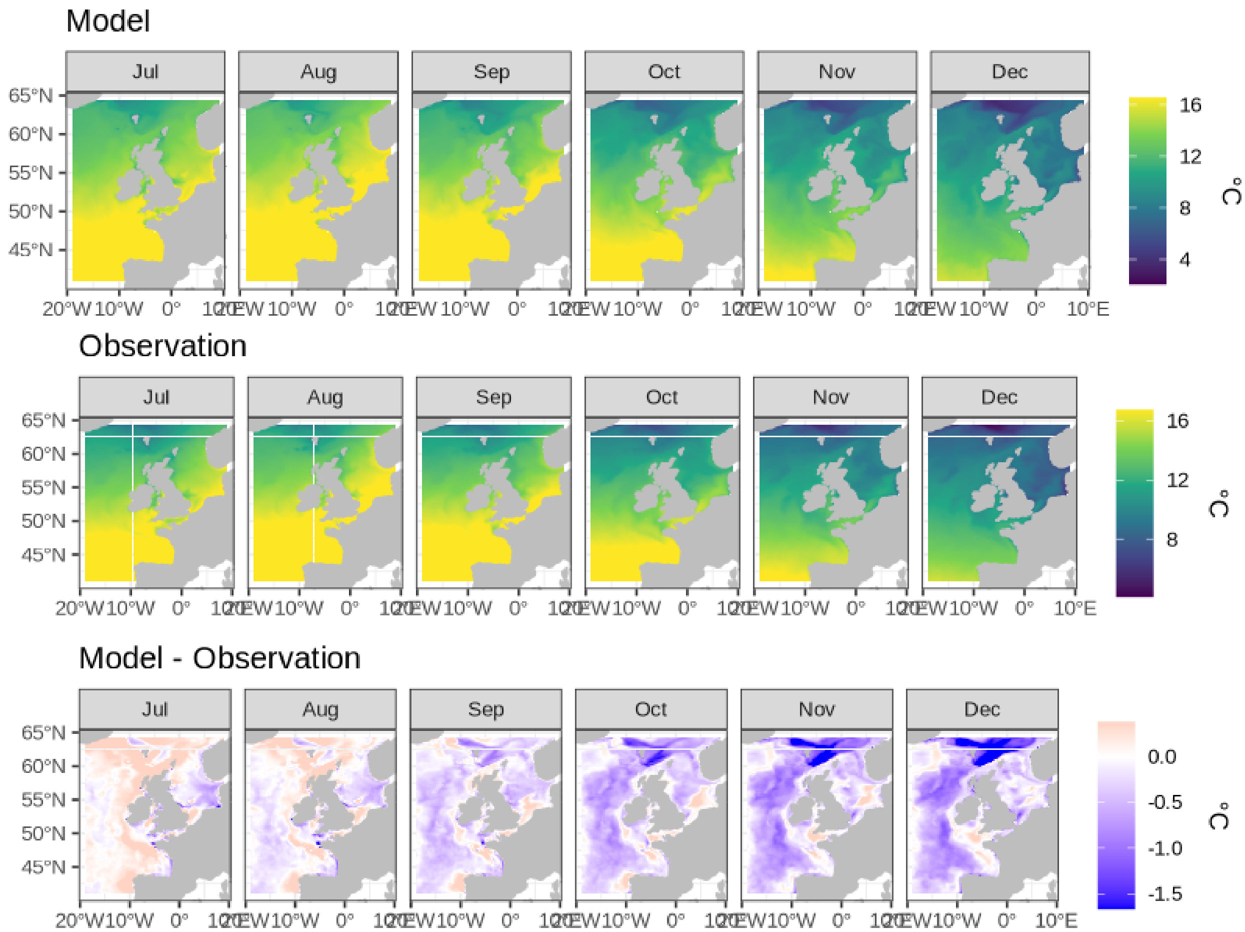


Figure 22.4: Monthly mean surface temperature for the model, observation and the difference between model and observations. For clarity, the maximum values are capped to the 98th percentiles.

22.4 Regional assessment of model performance for sea surface temperature

We assessed the regional performance of the model by comparing the model with observations in a number of regions. The regions considered are mapped below.

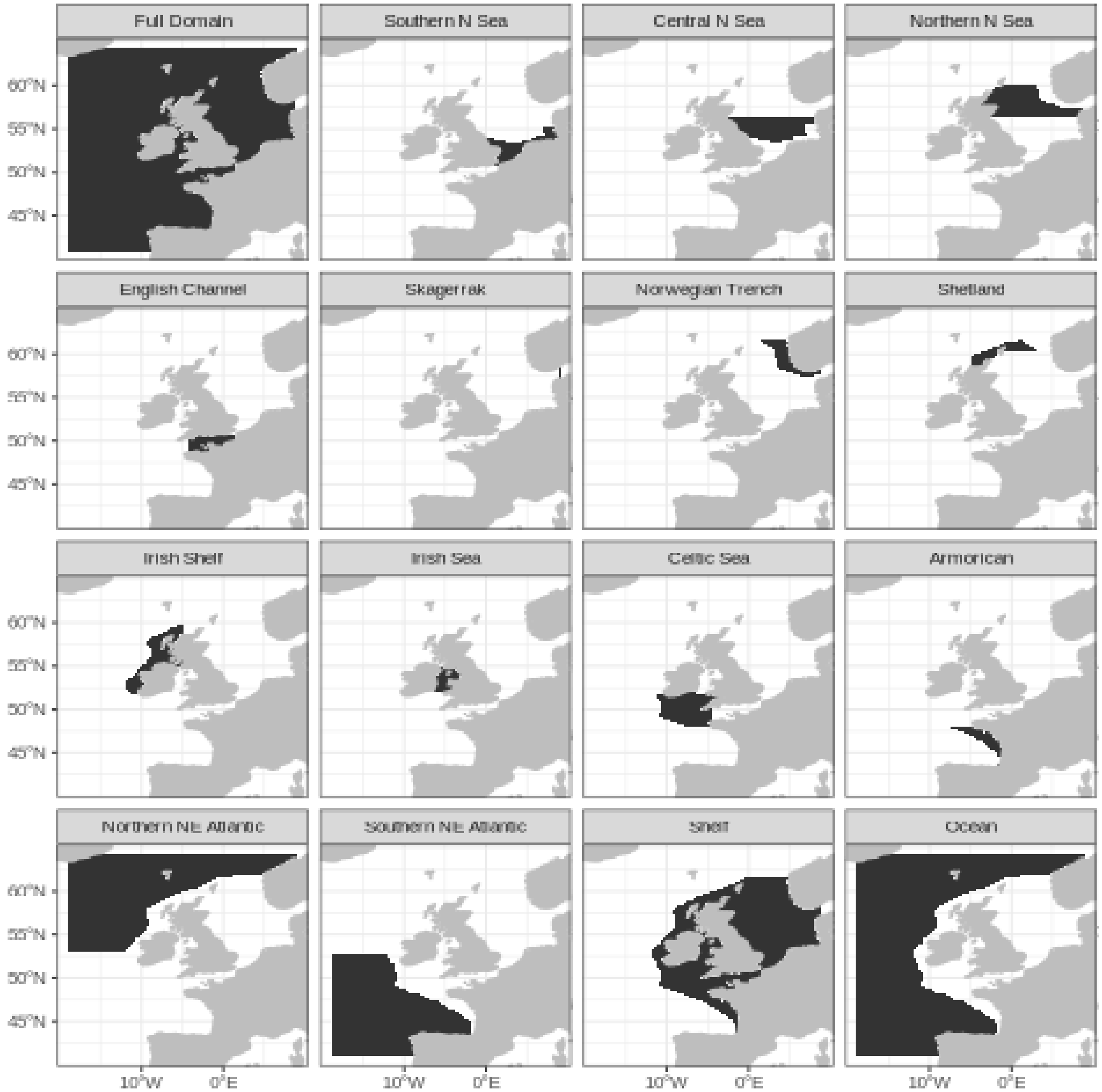


Figure 22.5: Regions used for validation of sea surface temperature.

Time series were constructed comparing the monthly mean of the spatial average sea surface temperature in each region. The spatial average was calculated using the mean of all grid cells within each region, accounting for grid cell area.

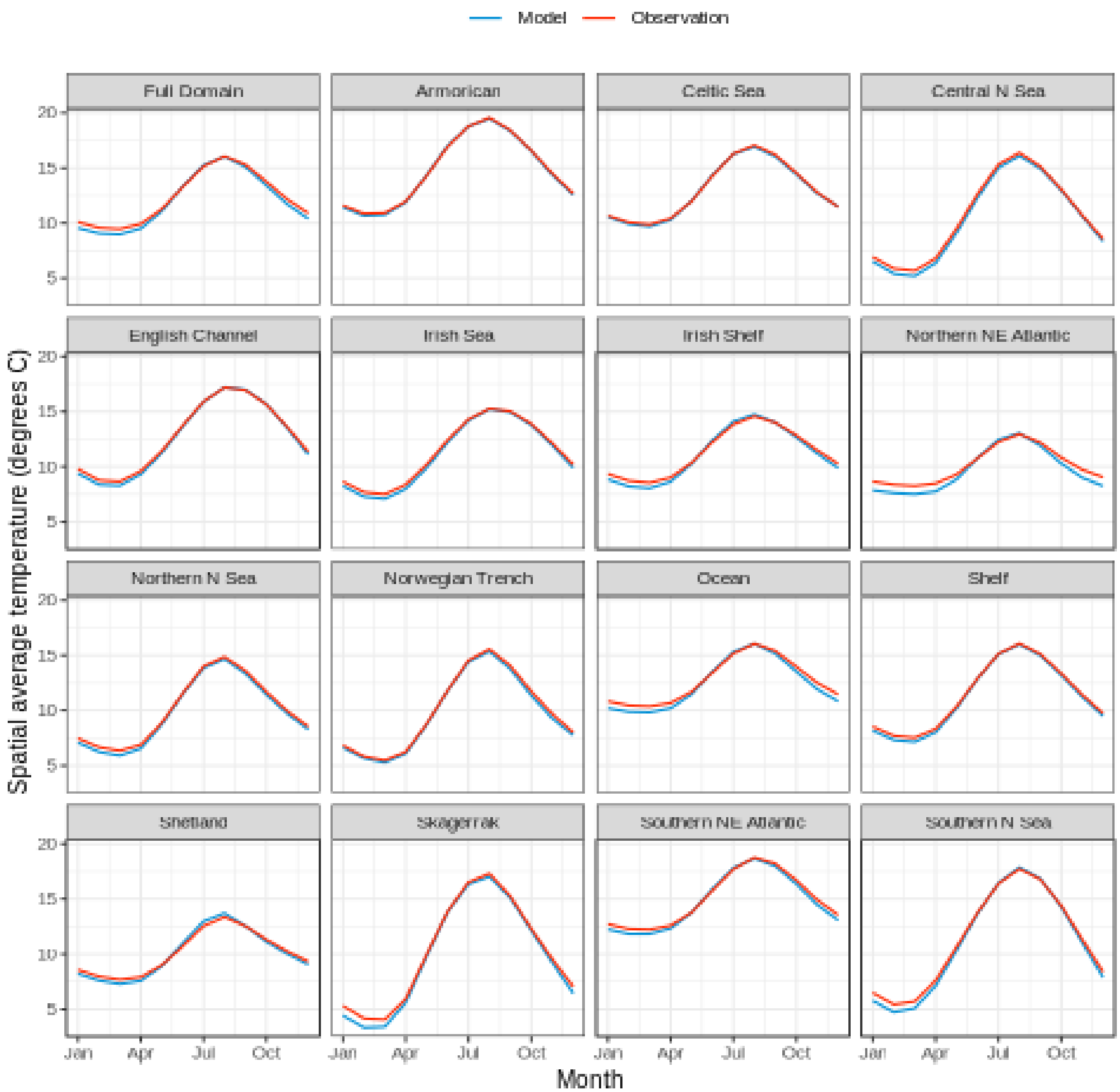


Figure 22.6: Seasonal cycle of sea surface temperature for model and observations for each region. The spatial average is taken over the region.

The table below shows the average bias of sea surface temperature in each region. The bias is calculated as the modelled value minus the observed value. A positive bias indicates that the model overestimates the observed value, while a negative bias indicates that the model underestimates the observed value.

Region	Temporal correlation	Bias	RMSD
Full Domain	1.00	-0.30	0.60
Armorican	1.00	-0.07	0.35
Celtic Sea	1.00	-0.08	0.30
Central North Sea	1.00	-0.30	0.42
Channel	1.00	-0.17	0.36
Irish Sea	1.00	-0.24	0.42
Irish Shelf	1.00	-0.19	0.45
Northern North East Atlantic	0.99	-0.46	0.84
Northern North Sea	1.00	-0.26	0.36
Norwegian Trench	1.00	-0.20	0.33
Ocean	1.00	-0.35	0.68
Shelf	1.00	-0.20	0.39
Shetland	1.00	-0.10	0.36
Skagerrak	1.00	-0.38	0.52
Southern North East Atlantic	1.00	-0.25	0.41
Southern North Sea	1.00	-0.30	0.54

Table 22.1: Summary of performance of the model sea surface temperature in each region. The bias (°C) column represents the spatial average of the annual mean modelled value minus the observed value. The temporal correlation column represents the spatial mean of the temporal correlation between the model and observations per grid cell.

22.5 Can the model reproduce spatial patterns of sea surface temperature?

The ability of the model to reproduce spatial patterns of sea surface temperature was assessed by comparing the modelled and observed temperature at each grid cell. We calculated the Pearson correlation coefficient between the modelled and observed temperature at each grid cell.

This was carried out monthly and using the annual mean in each grid cell

Time period	*r*
Annual mean	0.99
Jan	0.98
Feb	0.98
Mar	0.98
Apr	0.98
May	0.99
Jun	0.99
Jul	0.99
Aug	0.99
Sep	1.00
Oct	0.99
Nov	0.99
Dec	0.98

Table 22.2: Pearson correlation coefficient between modelled and observed sea surface temperature at each grid cell. The correlation was calculated monthly and using the annual mean in each grid cell.

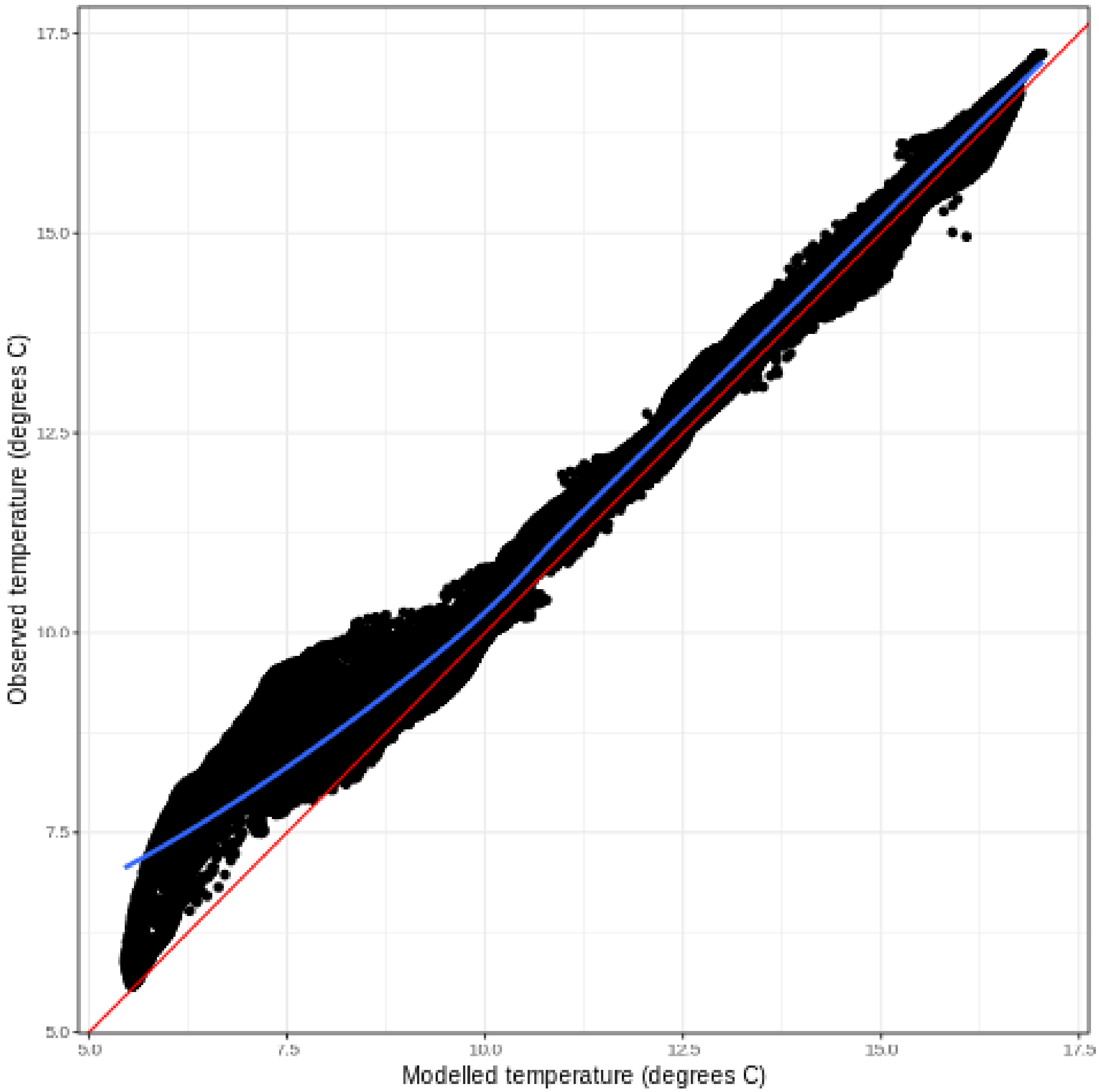


Figure 22.7: Scatter plot of modelled and observed annual average sea surface temperature in grid cells. The red line is the 1:1 line. The blue line is the linear regression of the modelled and observed values. The slope of the blue line is the slope of the regression line.

The ability of the model to reproduce multi-year trends in sea surface temperature was assessed by comparing the modelled and observed time series of annual temperature across each region.

The figure below shows the average temperature in each region

Scale for colour is already present. Adding another scale for colour, which will replace the existing scale.

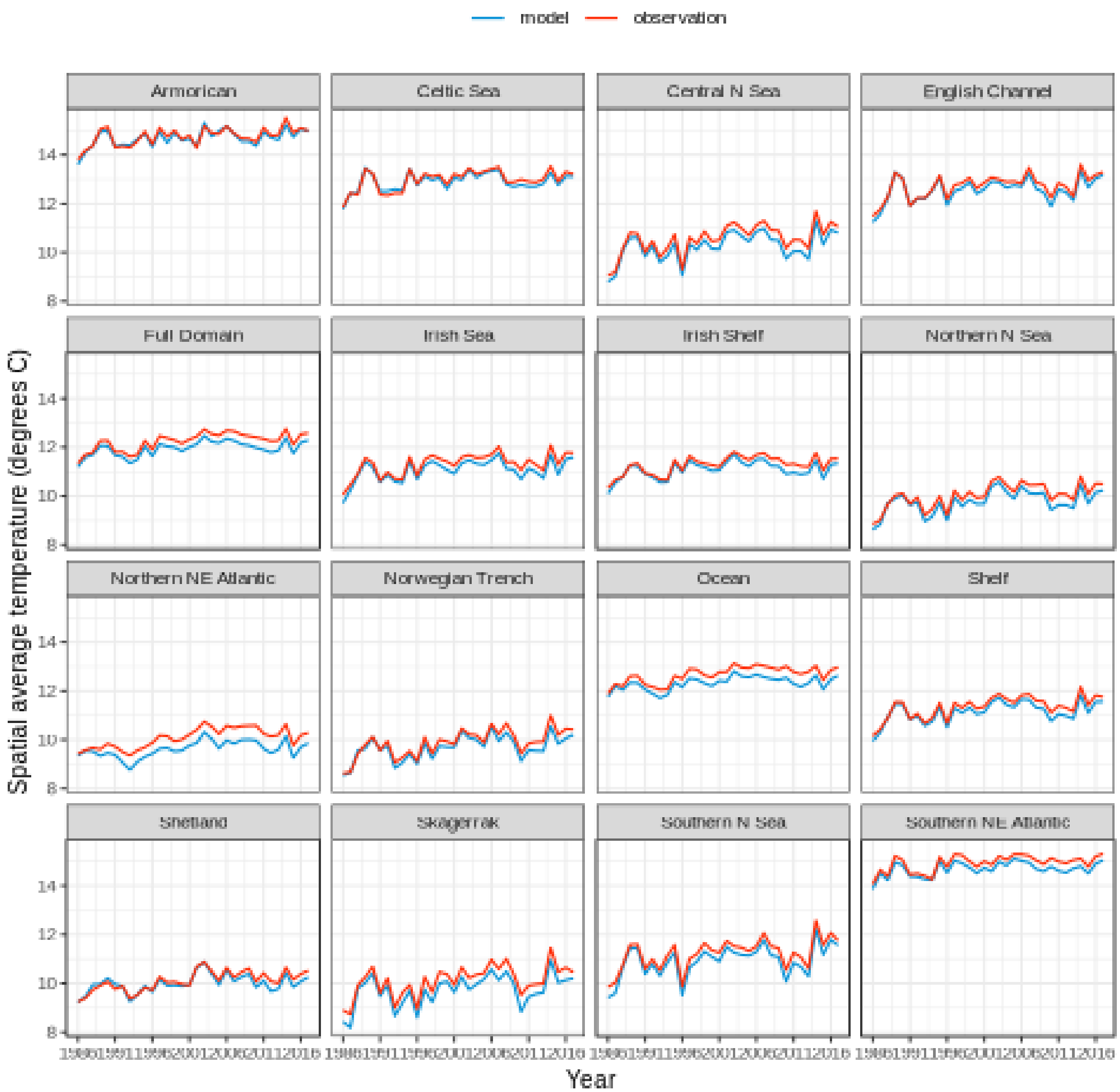


Figure 22.7: Changes in sea surface temperature for model and observations for each region for the period 1986-2017. The spatial average is taken over the region.

22.6 Data Sources for validation of temperature

Good, S.; Fiedler, E.; Mao, C.; Martin, M.J.; Maycock, A.; Reid, R.; Roberts-Jones, J.; Searle, T.; Waters, J.; While, J.; Worsfold, M. The Current Configuration of the OSTIA System for Operational Production of Foundation Sea Surface Temperature and Ice Concentration Analyses. Remote Sens. 2020, 12, 720, doi:10.3390/rs12040720

URL: https://data.marine.copernicus.eu/product/SST_GLO_SST_L4_REP_OBSERVATIONS_010_011/description

From 2022 onwards, the near-real time product is used: https://data.marine.copernicus.eu/product/SST_GLO_SST_L4_NRT_OBSERVATIONS_010_001/description