Improved understanding of nitrate trends, eutrophication indicators, and risk areas using machine learning

Deep S. Banerjee[1,2] **and Jozef Skákala**[1,2]

[1]Plymouth Marine Laboratory, PL1 3DH Plymouth, UK
[2]National Centre for Earth Observation, PL1 3DH Plymouth, UK

Correspondence: Deep S. Banerjee (dba@pml.ac.uk)

Abstract. Nitrate is an essential inorganic nutrient limiting phytoplankton growth in many marine environments. Eutrophication, often caused by nitrogen deposition, is a re-occurring problem in coastal regions, including the North-West European Shelf (NWES). Despite their importance, nitrate observations on the NWES are costly to obtain and thus sparse in both time and space. We demonstrate that machine learning (ML) can generate, from sparse observations, a skilled, gap-free, bi-decadal (1998–2020) surface nitrate dataset. We demonstrate that the effective resolution (scales on which the dataset is skilled) is slightly coarser than the 7 km and daily resolution of the product but still completely sufficient to analyse nitrate dynamics on a monthly scale. With such a dataset we are able to (i) highlight the coastal regions that show strong summer nutrient limitation, covering eutrophication problem areas identified by monitoring bodies (i.e. OSPAR), but also other regions, such as the southern Irish coastline and parts of the Irish Sea. Our results could indicate greater potential for eutrophication events in regions subject to high-riverine-nutrient-discharge scenarios. (ii) We demonstrate that bi-decadal 1998–2020 trends in coastal nitrate, responding to long-term policy-driven reduction in riverine discharge, are mostly modest, with a notable exception of the Bay of Biscay. (iii) We show that winter nitrate plays a relatively minor direct role in the intensity of the phytoplankton bloom the following spring, which can have some implications for using winter inorganic nitrogen as an indicator of eutrophication (as often included by OSPAR). The last two results are consistent with recent findings in the literature (Axe et al., 2022; Devlin et al., 2023; Van Leeuwen et al., 2023). We propose using the nitrate dataset for data assimilation and hypothesise that it has the potential to substantially improve phytoplankton forecasts in operational runs.

1 Introduction

Nitrogen is one of the most important components of organic matter, needed for primary production in relatively large concentrations, as demonstrated by the Redfield ratios (Tett et al., 1985). Despite its large abundance (the Earth's atmosphere comprises 78 % of nitrogen as N_2), it is non-trivial to obtain nitrogen in forms useful for plants. As a consequence, nitrogen is often the most limiting nutrient for plant or algae growth, including the coastal marine environment (Ryther and Dunstan, 1971; Board and Council, 2000). Nitrogen fixation, converting atmospheric nitrogen to forms useful for life, happens through various biotic and abiotic pathways, resulting in ammonium, nitrite, and nitrate (Noxon, 1976; Hill et al., 1980; Postgate, 1998; Beman et al., 2008; Voss et al., 2013). Nitrate in the ocean is the primary nutrient for phytoplankton, with phytoplankton uptake enabling nitrogen flows into higher trophic levels and various detrital and dissolved forms of organic matter. In a nitrogen-limited environment, excess nitrate concentrations, primarily originating from agricultural runoff and industrial wastewater discharge, can stimulate harmful eutrophication events (Withers et al., 2014; Nazari-Sharabian et al., 2018). The thick layer of algae produced by these events may reduce oxygen ventilation at the surface, and after the algae die off and sink, decomposers may consume vast amounts of oxygen, leading to marine hypoxia in the bottom part of the water column (Rabalais et al., 2002; Diaz and Rosenberg, 2008). Furthermore, eutrophica-

tion events are often dominated by species that produce toxins (leading to harmful algae blooms, HABs), which have detrimental effects on the marine ecosystem by causing fish kills, contaminating seafood, and even posing risks to human lives (Anderson et al., 2012). Additionally, high nitrate concentrations may lead to excessive production of organic matter under certain circumstances, which, upon decomposition, increases CO_2 concentrations, contributing to ocean acidification (Doney et al., 2009). Eutrophication is a fundamental problem in many shelf sea and coastal areas (Rabalais et al., 2009), where nitrate monitoring and prediction, along with other indicators (e.g. chlorophyll, dissolved oxygen, phytoplankton species), provide an essential tool informing marine management and policy.

An important region subject to eutrophication is the North-West European Shelf (NWES). The NWES is impacted by significant river inputs, such as the Elbe, Rhine, Loire, Seine, Scheldt, Meuse, Humber, Weser, and Thames, which introduce substantial freshwater and nutrients into the region, influencing salinity and water properties (Sonesten et al., 2022). Open ocean–shelf exchange, especially transport of nutrients and carbon across the shelf break, plays another vital role in the NWES ecosystem dynamics (Huthnance et al., 2009). The NWES has high ecological importance due to its high biological productivity, underpinning significant commercial fisheries and carbon sequestration (Pauly et al., 2002; Borges et al., 2006; Jahnke, 2010). Until the 1980s, the NWES, particularly near the German Bight and the Westerschelde Estuary, experienced notable shifts in nutrient distribution, primarily driven by increased continental nutrient inputs. Riverine discharges, particularly from the Rhine and Elbe, have been identified as major contributors to nutrient dynamics in the region (Brockmann and Eberlein, 1986; Radach, 1992), having adverse effects on the local ecosystem. However, EU regulations following the OSPAR Convention in 1992 have substantially decreased nitrate deposition into the NWES (Soetaert et al., 2006; Radach and Pätsch, 2007; Lenhart and Große, 2018; Burson et al., 2016; Axe et al., 2022; Sonesten et al., 2022).

The NWES nitrate concentrations are operationally simulated and predicted (Skákala et al., 2018); however, NWES nitrate observations are too sparse to properly constrain the simulated nitrate through data assimilation. Existing works have developed statistical algorithms to derive nitrate from satellite observations (Durairaj et al., 2015; Chen et al., 2023), but these have so far been developed either for the global open ocean or for regions other than the NWES (such as different regions in Asia or California, Yu et al., 2021; Chen et al., 2023), and those algorithms are unlikely to work for the NWES. The current operational NWES system is mainly constrained by much more robust satellite temperature and chlorophyll observations (Skákala et al., 2018, 2021, 2022) and avoids assimilating nutrients entirely. Furthermore, due to its univariate nature, the operational system fails to directly constrain most of the non-

assimilated variables, including nutrients. Consequently, the nitrate reanalyses and forecasts produced by the operational system are known to have substantial biases, inherited from the model free run (Skákala et al., 2018, 2022). Although the simulated physics and chlorophyll from the reanalysis are well validated against observations (Skákala et al., 2018, 2022), the nitrate NWES product is of more limited use.

In this work we develop and validate a new bi-decadal NWES nitrate product derived from available observations using advanced machine learning (ML) algorithms. The nitrate product is developed for the ocean surface, where nutrients have the potential to most significantly drive phytoplankton growth. This is, to our knowledge, by far the most complete and detailed observation-based sea surface nitrate dataset on the NWES. Unlike the NWES operational reanalysis, the dataset validates skilfully against the independent observations. Using our NWES nitrate product, we are able to discuss several important topics, like the impact of winter nitrate pre-conditioning on interannual phytoplankton variability, the most nutrient-limited coastal NWES geographic areas, and the trends in nitrate concentrations on the NWES. To do so, we maximise our reliance on the observational data and use ML and modelling to effectively fill the large data gaps, either through statistics or through dynamical consistency imposed by deterministic modelling.

2   Methodology

2.1   The ML model

We used as the ML model a feed-forward neural network (NN) designed through the Autokeras library using a structured data regressor (Jin et al., 2019). This approach streamlined the process of hyperparameter optimisation and model architecture discovery through an automated procedure, significantly reducing the need for manual intervention. The routine follows an iterative trial-and-error approach by looping through several possible combinations of various hyperparameters, such as learning rate, optimisers, and the number of dense layers with different combinations of nodes. It then selects the best model architecture with the highest skill score against the validation dataset and saves it for final prediction against the test data.

The final model architecture comprised several layers (see Fig. S1 in the Supplement): (i) the input layer, with 25 nodes, corresponds to the input features or predictors; (ii) a multi-category encoding layer, encoding categorical features (i.e. month and day of the year) into a numeric form that can be understood by the network; (iii) a normalisation layer, which normalises the input data to improve model training by facilitating improved model convergence time and better performance, avoiding dominance of features with larger magnitudes, and providing better stability; (iv) two dense layers

with 128 and 256 nodes, respectively, with each of these layers being followed by a rectified linear unit (ReLU) activation function to introduce non-linearity into the model; (v) a dropout layer that is applied to prevent overfitting by randomly dropping out a fraction of neurons during the training phase; and (vi) a regression head with a single node that produces the final prediction, i.e. nutrient concentration. During the training phase, the model's performance was evaluated using standard metrics, i.e. mean squared error (MSE), and the relative error was then estimated with respect to the validation dataset.

2.2 Data

2.2.1 The input features

To avoid biases towards operational models, the NN model input features were selected to be either observational data or reanalyses of variables closely constrained by the observations. One of the main challenges in developing ML for environmental applications is combining (often sparse) data with typically inconsistent domains of coverage across various spatial and temporal scales. The most robust NWES-focused observational datasets are obtained through satellite optical measurements. For physics, these are data such as sea surface temperature (SST) or altimetry. For biogeochemistry, the most typically derived dataset is surface chlorophyll a concentration obtained from the ocean colour (OC). Recently, new remote-sensing algorithms have been developed to partition the total chlorophyll concentration into phytoplankton functional types (PFTs) largely based on phytoplankton size classes (Brewin et al., 2010, 2017), and PFT chlorophyll products are now operationally assimilated into the NWES model (Skákala et al., 2018). The reanalyses from the NWES operational model have been found to have very close match-ups with the assimilated observations (Skákala et al., 2018, 2022) and act as natural extensions of the observations, forming a complete dataset on a gridded domain. We have extracted a range of NN model features from NWES physical–biogeochemistry reanalysis products, namely NWSHELF_MULTIYEAR_BGC_004_011 and NWSHELF_MULTIYEAR-_PHY-_004_009 (downloadable from the EU Copernicus portal, https://data.marine.copernicus.eu/product/NWSHELF_ MULTIYEAR_BGC_004_011/services, last access: 20 June 2025; see also Kay et al., 2016), covering the 1998–2020 period with daily and 7 km spatial resolution. The product is based on assimilating satellite SST, temperature, and salinity profiles, as well as OC PFT chlorophyll, into the operational Nucleus for European Modelling of the Ocean (NEMO; Madec et al., 2017) model coupled through the Framework for Aquatic Biogeochemical Models (FABM; Bruggeman and Bolding, 2014) to the biogeochemistry European Regional Seas Ecosystem Model (ERSEM; Baretta et al., 1995; Butenschön et al., 2016). The extracted

features were for (i) SST; (ii) chlorophyll ocean surface concentrations from four PFTs, which were assimilated into NEMO-FABM-ERSEM (diatoms, microphytoplankton, nanophytoplankton, and picophytoplankton); (iii) total surface phytoplankton carbon; (iv) total surface chlorophyll; and (v) total surface net primary production. Although these features were selected from the reanalysis, they correspond to either the assimilated variables (SST, PFT chlorophyll) or the variables which are dynamically very close to the assimilated PFT chlorophyll (phytoplankton carbon, net primary production) and therefore well constrained by the assimilation. Additionally, we also included SST observations from the global ocean OSTIA product (Good et al., 2020; Donlon et al., 2012) in the input feature dataset. Interestingly, our tests (not shown here) indicated that the NN model did not perform as well without this additional SST dataset, so both sources of SST information (reanalysis and OSTIA) were used.

We have also used input features describing riverine discharge into the ocean. These included riverine discharge data for oxygen and nutrient loads (i.e. nitrate, phosphate, silicate, ammonia, and oxygen) at all relevant river mouths in the NWES domain. Daily time series of river discharge are used for 1998–2017. From 2018, only climatologies were available and were used for the remaining 2018–2020 period. The daily riverine data were obtained from an updated version of the river dataset from Lenhart et al. (2010). The climatology of daily discharge data was taken from the Global River Discharge Database and the Centre for Ecology & Hydrology (Young and Holt, 2007). Unlike a hydrodynamic model, an NN model does not follow any advection mechanism or transport to carry the effect of river discharge over space and time. To account for advection, the NN model would have to ideally include time-lagged riverine inputs, where the time lag would increase with the spatial distance from the river mouth. This would hugely increase the complexity of the NN model. In this work we have decided to opt for less complex models and avoid using time-lagged NN inputs. To consider a nearly instantaneous riverine effect in a simplified way, we have distributed the riverine discharge data around the river discharge point sources. This was done by spatially extrapolating all the runoff variables at all the daily discharge points over a 50 km circular radius surrounding the main discharge point, by making the inputs decay inversely from the maximum at the centre to a zero value at the edge (see Fig. 1). However, such a scheme is clearly a major simplification of the real river impact (e.g. Painting et al., 2013; Lenhart and Große, 2018) and should ideally be improved upon in future work.

Another dataset that was used as input features into the NN model was the ERA5 atmospheric reanalysis (Hersbach et al., 2020), which has a horizontal resolution of 0.25°. The input feature dataset includes variables such as downwelling shortwave radiation at the ocean surface, specific humidity, temperature at 2 m above the ocean surface, sea level pres-
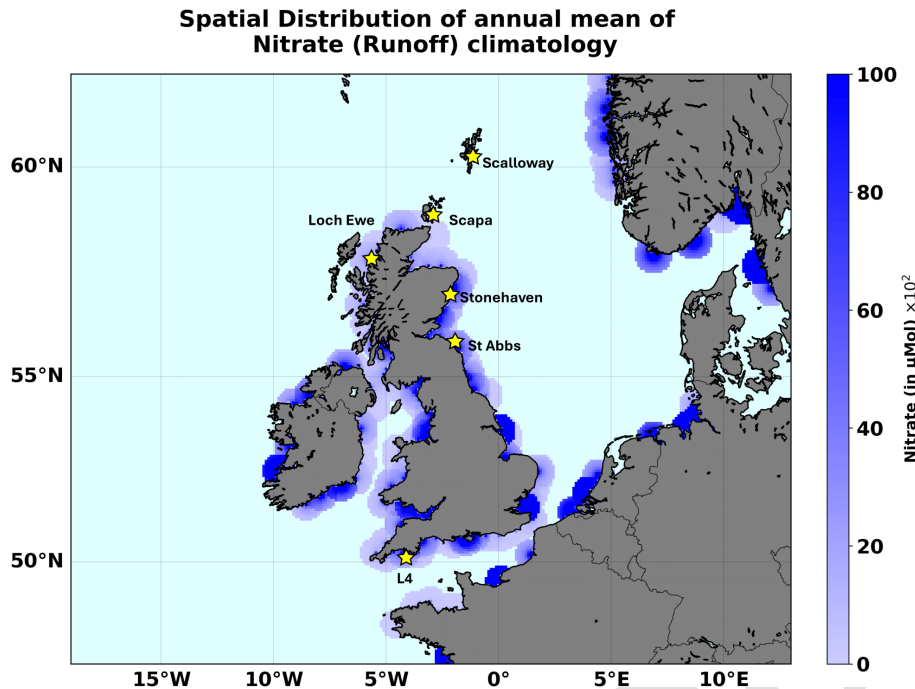
Figure 1. The areas with river input on the NWES domain. We mark the locations of the stations providing the test data for this study: L4 and the five Scottish coastal stations.

sure, total precipitation, and zonal and meridional wind components at 10 m above the ocean surface. These near-surface atmospheric drivers play a crucial role in governing and redistributing surface nitrate through air–sea interactions and atmospheric deposition of nitrogen, which accounts for one-third of the non-recycled nitrogen supply in the ocean and up to around 3 % of annual new biological production (Arrigo et al., 2008). The atmospheric data, however, did not include any direct products for atmospheric nitrogen deposition. Finally, we used structural temporal and geographic data, i.e. the time of the year (month, day), latitude, longitude, depth, and bathymetry, as input features. This type of information enables the model to learn the geographic patterns in nitrate, including its seasonal climatology, substantially enhancing its predictive accuracy.

All the input features were considered at the same time as the predicted nitrate. The input features' importance was ranked in the SHapley Additive exPlanations (SHAP) analysis, presented in Fig. S2 in the Supplement. The SHAP analysis indicates that the structural input features are among the most important, followed by some atmospheric and oceanic physics features (incoming shortwave radiation, SST), which can partly account for seasonal climatology as well. Specific riverine discharge input features (e.g. of nitrate itself) are highly important too (Fig. S2), with a range of biogeochemical variables (total surface chlorophyll, total surface net primary production, and total surface carbon) being around the middle of the importance ranking. As already mentioned, we also tested versions of the NN model with a reduced number of input features (e.g. removing the less important features from the SHAP analysis), but the model's performance became slightly worse compared to that of the previous model.

2.2.2 The predicted nitrate data and the training and validation process

The 1998–2018 nitrate observations were obtained from the International Council for the Exploration of the Sea (ICES) data portal (https://www.ices.dk, last access: 20 June 2025). The extracted dataset spans a geographical range from 19° W to 10° E in longitude and 47 to 62° N in latitude, ensuring a broad representation of the dynamics and variability of the NWES region. The ICES data were obtained from a wide range of in situ measurements, e.g. by cruises, floats, moorings, and buoys. In the training and validation process, the ML model inputs were linearly interpolated onto the ICES data locations, and then the ICES data from the 1998–2015 period, containing 43 572 relevant data points, were used for training and validation of the NN model (with 80 % of data used for training and 20 % for validation). Finally, the 2016–2018 ICES data, containing 2984 data points, were used as test data. The spatial coverage of all the ICES training, validation, and test data is shown in Fig. 2.

Several other observations were used as test data to demonstrate the ML model's skill: (i) nitrate data from the L4 station, which is operated by the Western Channel Observatory (https://www.westernchannelobservatory.org.uk/data.php, last access: 20 June 2025) and is located in the western
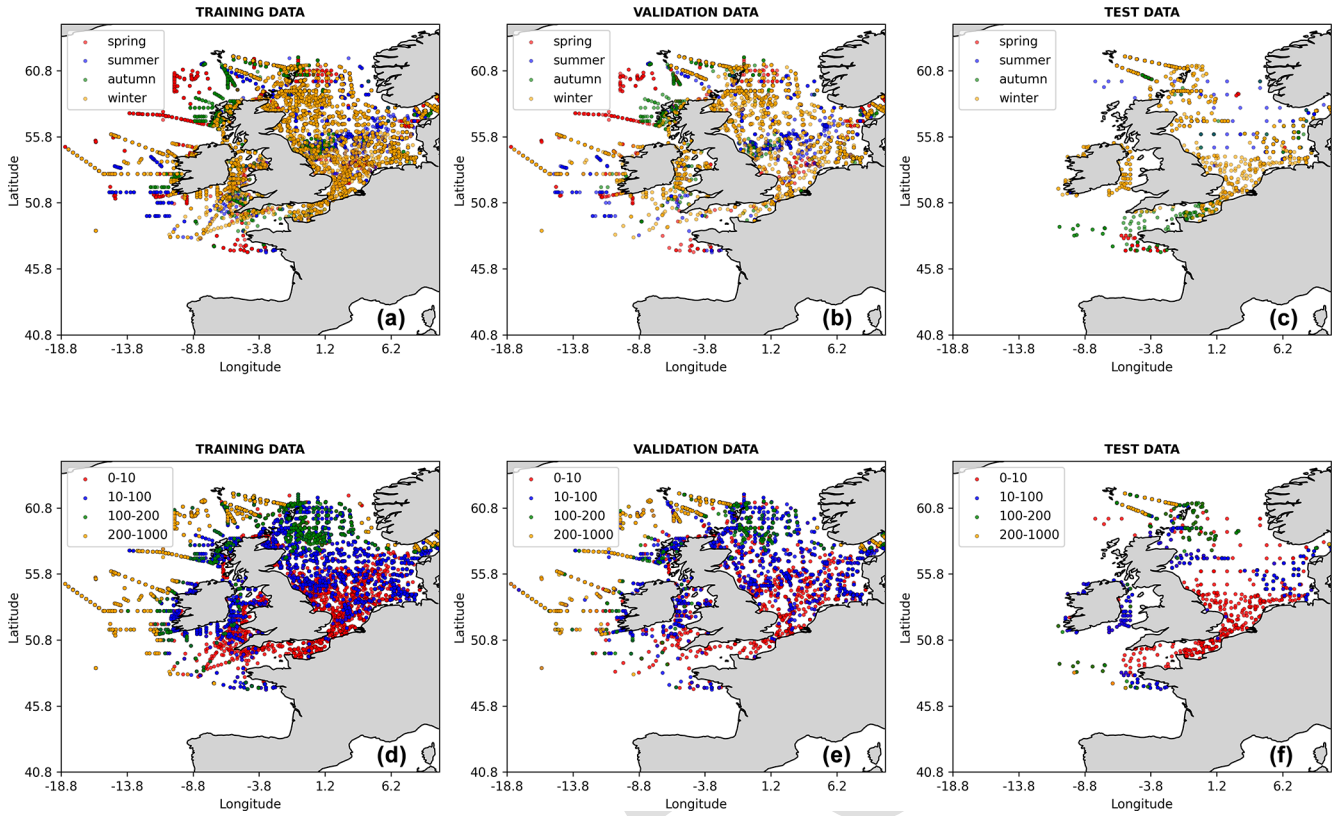
Figure 2. The locations of the ICES data used in this study, split into training, validation, and test data. The data points are coloured with respect to the season in which the measurement was taken **(a–c)** and the depth range (in m) at which the measurement was taken **(d–f)**.

English Channel approximately 13 km from the Plymouth Sound, providing one of the longest continuous ecological time series in the world (Harris, 2010). The L4 nitrate dataset covered the whole 1998–2020 period, and despite several data gaps, during most of this period it was sampled with a frequency of approximately 5–7 d. (ii) Another independent nitrate test dataset was obtained from the Scottish Coastal Observatory Dataset (https://data.marine.gov.scot/dataset/scottish-coastal-observatory-dataset-1997-2020, last access: 20 June 2025; see also Bresnan et al., 2016; Hindson et al., 2018), covering five locations near the coastline of Scotland (Loch Ewe, Scalloway, Scapa, St Abbs, Stonehaven) and providing time series of differing lengths: from the longest, covering the 2008–2020 period (Scapa, Scalloway), to the shortest, covering the 2017–2020 period (St Abbs). The measurements at these stations typically had a frequency of 5–7 d. The locations of the L4 station and the five Scottish stations are all marked in Fig. 1. It should be noted that neither the L4 data nor the data from the Scottish stations were included in the ICES dataset.

Finally, after validating the NN model, we ran it for the full 1998–2020 period across the whole Copernicus NWES reanalysis domain (see Fig. 3), using Copernicus reanalysis, river input, and ERA5 atmospheric forcing input (with both inputs interpolated onto the Copernicus reanalysis domain)

to produce a gap-free, bi-decadal, daily 7 km resolution reconstruction of nitrate. This final dataset underpins the results from this study.

2.3 Skill metrics

We used the relative performance (RP) skill metric to compare the performance of the NN model from this study with the reanalysis:

$$\mathrm{RP(NN, Rean)} = 100 \cdot \frac{(|\mathrm{NN} - \mathrm{Obs}| - |\mathrm{Rean} - \mathrm{Obs}|)}{\mathrm{Obs}}. \quad (1)$$

In Eq. (1), "Obs" stands for observations, "NN" for predicted values by the NN model, and "Rean" for reanalysis. Negative values of the RP metrics from Eq. (1) indicate that the NN model outperforms reanalysis and vice versa.

The bias between any model, "Mod" ("Mod" could be either NN or reanalysis), and observations is defined as

$$\mathrm{Bias(Mod, Obs)} = \langle \mathrm{Mod} - \mathrm{Obs} \rangle, \quad (2)$$

where the averaging $\langle \ldots \rangle$ is taken through all the available matching model and observational data.

The bias-corrected RMSE (BC-RMSE) is defined as the RMSE after the bias h has been subtracted from the model:

$$\mathrm{BC\text{-}RMSE} = \sqrt{\langle (\mathrm{Mod} - \mathrm{Obs} - \mathrm{Bias})^2 \rangle}. \quad (3)$$
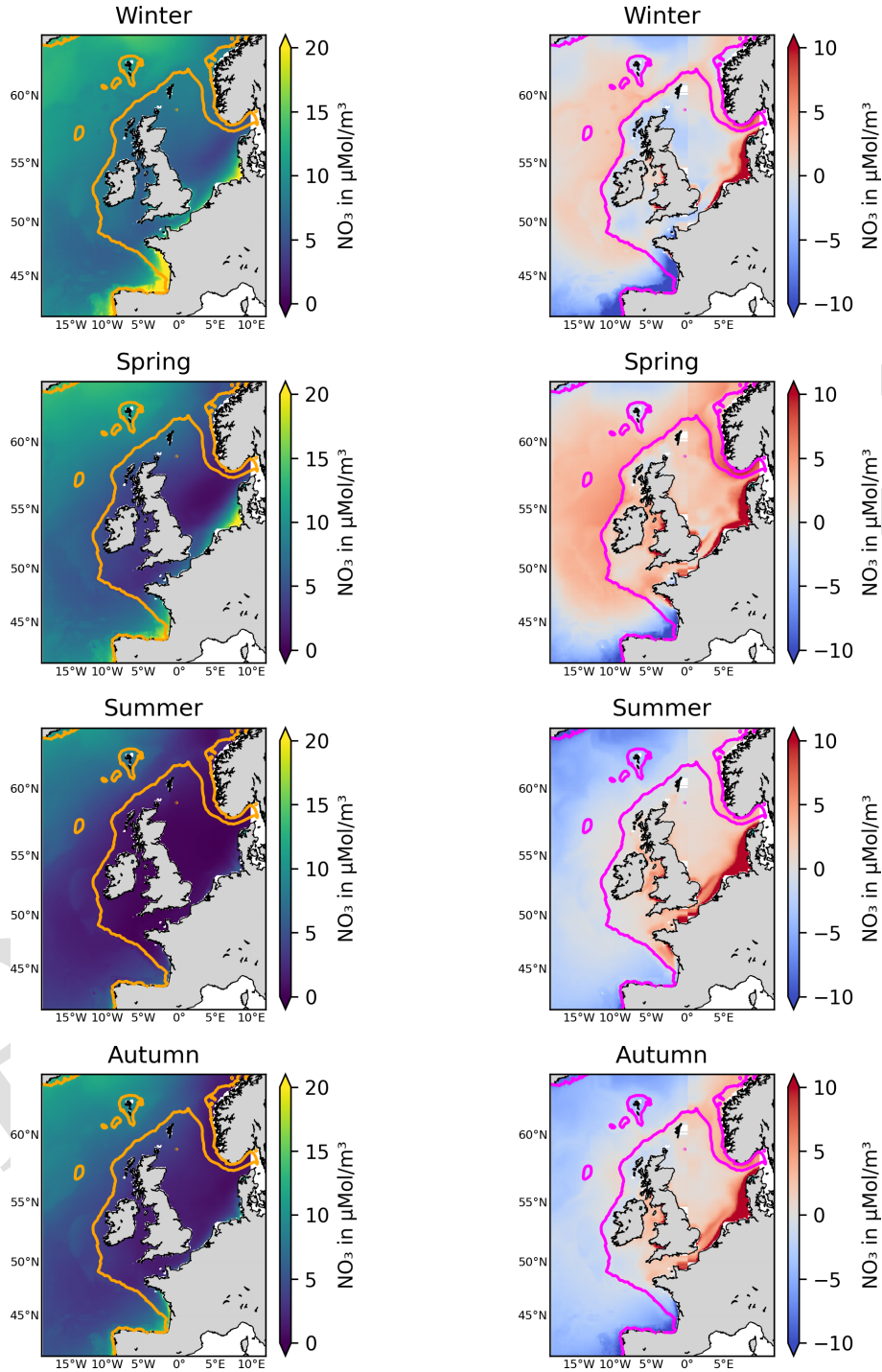
Figure 3. The left-hand panels show the NN-reconstructed 1998–2020 average surface nitrate concentrations for different seasons (corresponding to the dominant mode of temporal variability in nitrate). The right-hand panels show the same averages for the relative bias of the Copernicus surface nitrate reanalysis (Kay et al., 2016) with respect to the NN-reconstructed dataset (reanalysis minus NN-reconstructed dataset). The contours mark the NWES (bathymetry < 200 m).
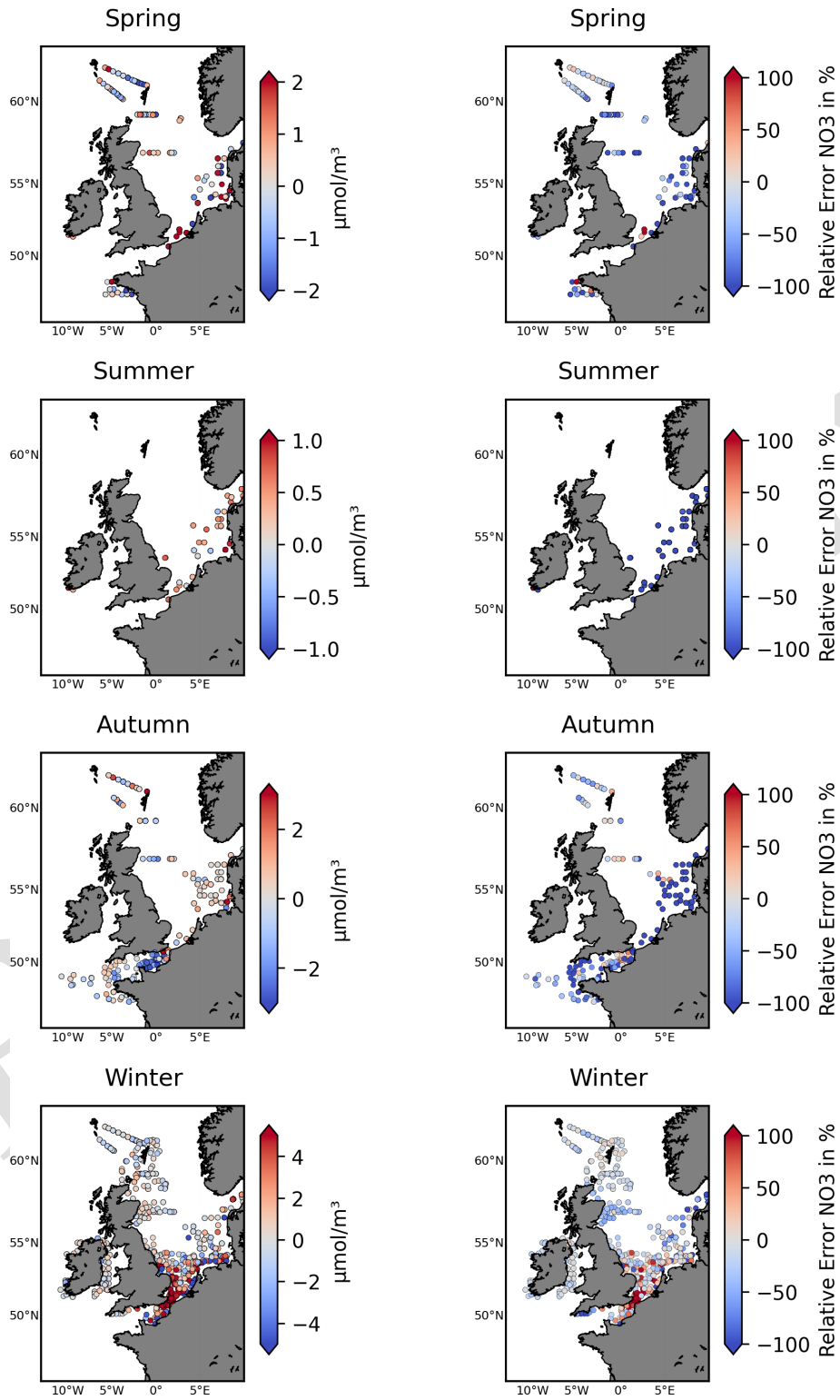
Figure 4. The left-hand panels show the ICES nitrate test data locations for different seasons, with the colour bar accounting for the NN model's skill (difference between predicted and observed nitrate: predicted−observed). The right-hand panels show the relative performance metrics comparing the NN model to the Copernicus reanalysis skill, as defined in Eq. (1). It marks the NN model's improvement (blue) or degradation (red) relative to the reanalysis when compared (in %) to the observed nitrate concentrations.

Table 1. The skill of the NN model in predicting nitrate compared with the Copernicus reanalysis (Kay et al., 2016). Skill is measured by bias (Eq. 2, in $\mu mol\,m^{-3}$), the bias-corrected root mean squared error (BC-RMSE; Eq. 3, in $\mu mol\,m^{-3}$), and Pearson's correlation (R). The rows represent different test data from ICES and the coastal stations. The last two rows show the skill of the NN model and the reanalysis in predicting interannual, low-pass-filtered time series (for details, see Fig. S3). For the five Scottish stations, we show only the averaged result of all the stations.

Test data	NN predicted			Reanalysis		
	Bias	BC-RMSE	R	Bias	BC-RMSE	R
ICES	0.62	2.37	0.72	3.18	6.15	0.27
L4	−1.03	1.85	0.79	0.20	2.22	0.72
Scalloway	0.98	2.1	0.8	1.08	2.76	0.64
St Abbs	−0.25	1.25	0.9	0.73	1.87	0.85
Scapa	1.52	1.52	0.86	0.33	2.13	0.69
Stonehaven	−0.53	0.97	0.95	0.09	2.04	0.78
Loch Ewe	1.57	0.92	0.93	0.57	1.1	0.89
L4 interannual	–	0.72	0.52	–	0.99	0.08
Scottish interannual	–	0.512	−0.144	–	0.756	0.204

Apart from the metrics from Eqs. (1)–(3), we also used Pearson correlation. Our tests show that the effective temporal resolution of the NN model (timescale on which it performed best relative to the test data) is around 15 d (the tests are not shown here, but some insight is provided by Fig. S3 in the Supplement). We therefore low-pass filtered all the compared data on a 15 d scale before any of the metrics from this section were applied.

3   Results and discussion

3.1   Model validation

Figure 4, Table 1, and Figs. S4–S5 in the Supplement demonstrate that the NN model shows very good skill relative to the test data from ICES, L4, and the Scottish stations and substantially outperforms the existing Copernicus reanalysis product for NWES nitrate. For example, the bias measured by the ICES test data has been reduced by more than 80 % relative to the reanalysis, the BC-RMSE has been reduced by more than 60 %, and the Pearson correlation has increased from 0.27 to 0.72. Comparison to the data from coastal stations shows less consistent improvement in terms of bias relative to the reanalysis, but the NN model still outperforms the reanalysis in BC-RMSE and R at each of the locations.

Because the nitrate time series are dominated by the seasonal signal, it is important to explore whether the model skill extends beyond predicting the local nitrate seasonal (e.g. monthly) climatology. This is much harder to validate, as one needs long-term time series at specific locations, which are rare. We have looked at the data from the L4 station and five Scottish locations to analyse the ML model's skill in capturing interannual variability of nitrate. The results (shown in Table 1 and Fig. S3) are more mixed: at the L4 station, which has the longest time record and richest dataset out of

all locations, the ML model performs very well in predicting the interannual nitrate time series. It is interesting that at the same location the reanalysis does a very poor job in doing the same (Table 1). At the Scottish stations, the ML model correctly captures the size of the interannual variability in nitrate, but it struggles to capture the variability itself (the R metrics in Table 1). It is however noteworthy that some of the time series at the Scottish locations are relatively short (see Sect. 2.1) and therefore not the most suitable for this type of analysis.

Finally, the test data selected from the ICES dataset are time separated from the training and validation data but are spatially located in largely overlapping regions (see Fig. 2). It is therefore important to explore the possibility that, due to geographic proximity, some ML skill has been transferred from the training and validation data to the test data. This is done in Fig. S6 in the Supplement, showing how the skill evolves as a function of spatial separation between the test data and the training and validation data. Although there is large variability in the skill, Fig. S6 shows no significant trend with spatial distance, indicating that the ML model's skill does not decrease (it even improves slightly) with the increase in spatial separation.

3.2   The bi-decadal nitrate product, trends, variability, and implications

Figure 3 shows the 1998–2020 seasonally averaged NWES nitrate concentrations. It is clear from the spatial nitrate distributions that the NN model does not capture the ∼ 7 km scale variability sufficiently, including the exact NWES boundaries, but it does capture coarse-resolution nitrate distributions reasonably well (see Fig. 5 for comparison with the World Ocean Atlas (WOA) product of Garcia et al., 2019). Similarly, our analyses (including Figs. S3–S5) suggest that the effective temporal resolution of the NN product is ∼15 d,

rather than daily. Figure 3 also provides seasonal comparison with the Copernicus reanalysis product, evaluating the significant reanalysis biases throughout the 1998–2020 period. The Copernicus reanalysis validation gives similar results to the validation from Kay et al. (2016), who compared the reanalysis with the North Sea Biogeochemical Climatology for 1960–2014 (Hinrichs et al., 2017).

Winter nitrate concentrations play an important role in the pre-conditioning of the spring bloom, which largely drives the NWES's biogeochemical seasonal cycles (Huisman et al., 1999; He et al., 2011). Winter dissolved inorganic nitrogen is used by OSPAR (in combination with other parameters) in its common procedure (OSPAR, 2005) as an important indicator of NWES eutrophication and the next season's growth (Axe et al., 2017; Topcu and Brockmann, 2021). The hypothesis that the intensity of the spring phytoplankton bloom is directly related to the abundance of nutrients in the winter before the bloom has been investigated here through nitrate.

In Fig. 6a, we find only limited evidence of the relationship between the winter (December–February) nitrate and spring bloom intensity; i.e. a statistically significant positive Pearson correlation has been found only in the western English Channel region, near the shelf break in the Celtic Sea, around the Bay of Biscay, and in the south-west of the model domain (accounting for at most 30 %–35 % of explained variance). Figure 6a also shows that these are regions where the interannual nitrate variability appears to be relatively large (> 5 % of the winter average in the majority of the region, with 10 %–20 % in specific sections, Fig. 6b) and therefore capable of revealing a stronger relationship with spring chlorophyll. For most of the domain, there is a lack of clear correlation between interannual winter nitrate and spring chlorophyll, which could be explained by the fact that both are driven by the interannual variability in the atmosphere (Dutkiewicz et al., 2001; Follows and Dutkiewicz, 2001; Ueyama and Monger, 2005; Henson et al., 2006; Zhai et al., 2013). Increased winds can lead to more mixing and elevated surface nutrients, whilst dampening blooms by transporting phytoplankton below the Sverdrup critical depth, as proposed by popular hypotheses explaining the North Atlantic spring blooms (Sverdrup, 1953; Huisman et al., 1999). Furthermore, there is a lack of complete agreement on what the dominant drivers of the spring bloom in the North Atlantic are, and arguments have been raised that support the view that blooms result much more from internal ecosystem dynamics (e.g. zooplankton control over phytoplankton, Behrenfeld and Boss, 2014) compared to what was assumed by traditional hypotheses focusing on physics. It is noteworthy that the weak link between winter nitrate and phytoplankton growth reported here is also consistent with recent results by Van Leeuwen et al. (2023).

In Fig. 6c we look at correlations between interannual time series of summer nitrate and chlorophyll concentrations. The high positive correlation in Fig. 6c indicates regions where growth is either strongly nitrate limited or limited by other drivers that are positively correlated with nitrate. Realistically, only other nutrients are likely to act as such drivers, i.e. phosphate (there is substantial phosphate limitation on the NWES, Skogen et al., 2004; Philippart et al., 2007; Loebl et al., 2009; Lenhart and Große, 2018; Burson et al., 2016; Grosse et al., 2017), so we can conclude that those regions are strongly nutrient limited in the summer (with the nutrient likely being nitrate). The coastal part of the regions with strong positive correlation delimits areas which could be more sensitive than others to high river nutrient loads. This does not automatically imply high risk of eutrophication, as this would also depend on other factors, such as the overall river outflow in each area and the socio-economic activity, but it indicates certain increased vulnerability. Figure 6c shows high summer nitrate and positive chlorophyll correlations, mostly in the southern North Sea region, the western English Channel, the Bay of Biscay, and the south-west of the domain. These are again the regions where the interannual fluctuations in summer nitrate are relatively large (Fig. 6d). Interestingly, the eutrophication problem areas, as identified by the OSPAR NWES eutrophication status reports (such as the south-eastern North Sea, coastal areas around Brittany, Axe et al., 2017; Devlin et al., 2023), fall under these vulnerable zones delimited in Fig. 6c. However, Fig. 6c also includes other regions, such as the eastern coastline of Scotland, the southern coast of Ireland, and zones in the Irish Sea.

Finally Fig. 7 shows 1998–2020 trends in winter nitrate concentration over the NWES domain. In most of the domain, no statistically significant nitrate trends have been detected, but some small negative trends ($\sim 0.02\,\mu\mathrm{mol\,m^{-3}\,yr^{-1}}$) were found in the southern North Sea and the north-east region near the Norwegian trench. Somewhat larger ($\sim 0.08\,\mu\mathrm{mol\,m^{-3}\,yr^{-1}}$) statistically significant negative trends have been found in specific locations of the Bay of Biscay. These results (e.g. from the southern North Sea) are broadly consistent with what has been reported for this period in the recent OSPAR report (e.g. Axe et al., 2017, 2022). These small trends may follow the smaller rates of reduction in the nitrate riverine inputs during the data period (1998–2020) compared to their large reduction in the 1980s and earlier 1990s (Duarte, 2009; Brockmann et al., 2018; Greenwood et al., 2019). However, it should be noted that significant reduction in atmospheric nitrogen input has been reported in the last few decades (Devlin et al., 2023).

4 Conclusions

In this work we have demonstrated that, using sparse observations across the North-West European Shelf (NWES), machine learning (ML) can be a powerful tool to reconstruct a spatially complete sea surface nitrate dataset over a period of 23 years. We have shown that the dataset has substantially better match-ups with independent test data than the existing
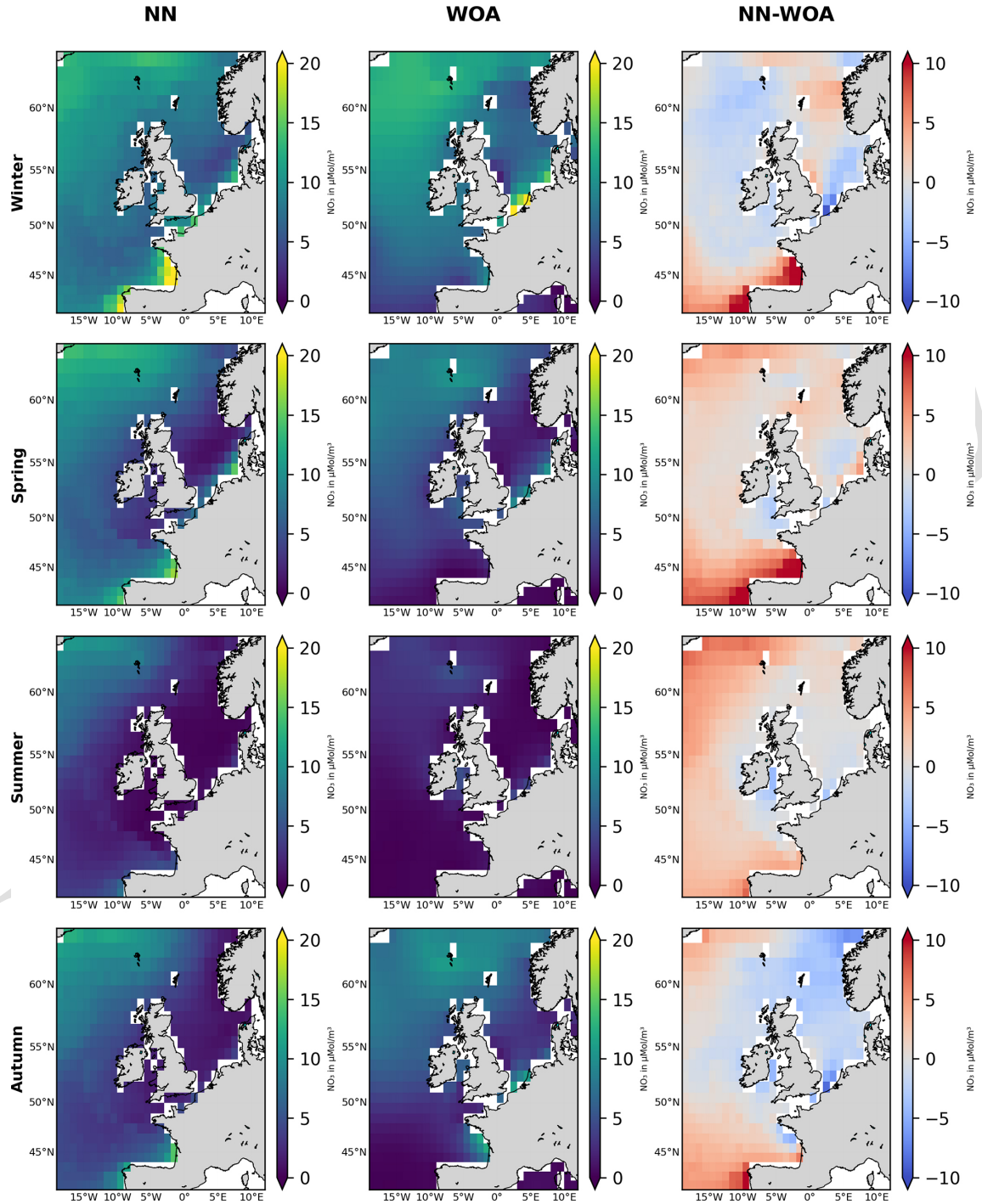
Figure 5. Seasonal comparison of the NN-predicted 1998–2020 surface nitrate averages (left-hand panels) with those of the World Ocean Atlas (WOA) 2018 product (middle panels, Garcia et al., 2019). Right-hand panels show the difference (NN-predicted averages minus WOA product averages) for corresponding seasons. The NN-predicted data are coarse grained on the WOA 0.25° spatial resolution scale. The focus is on spatial nitrate features, rather than nitrate concentration values, as the WOA averages data for the whole 1900–2018 period, including highly eutrophic decades in the 1970s and 1980s.
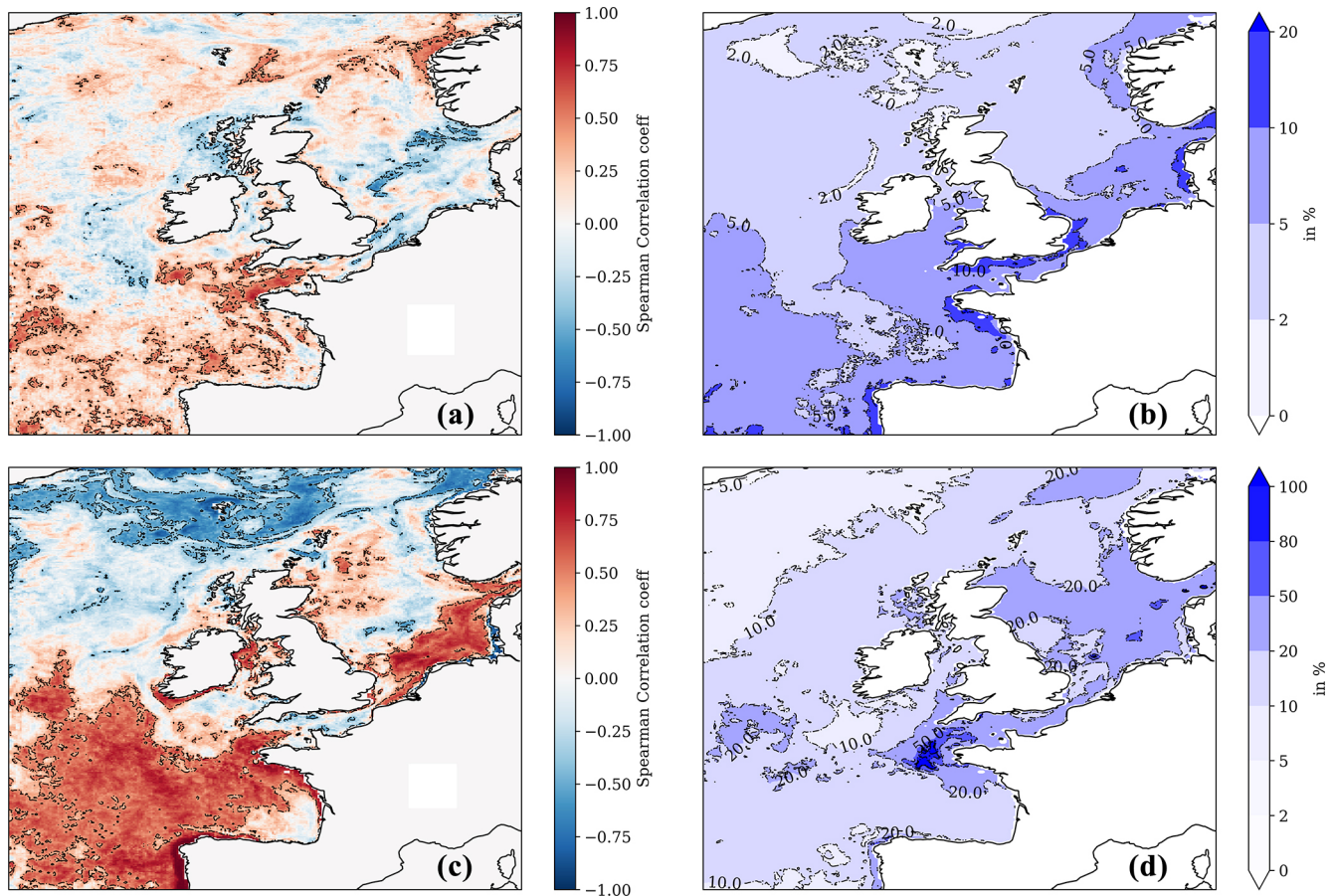
Figure 6. The upper left-hand panel **(a)** shows the Pearson correlation between the NN-predicted mean winter surface nitrate concentrations and the mean (following) spring surface chlorophyll concentration from the Copernicus reanalysis. The upper right-hand panel **(b)** shows the interannual variability for NN-predicted winter surface nitrate concentration (for 1998–2020, measured by the standard deviation) relative to the 1998–2020 winter mean (in %). The bottom-left panel **(c)** is similar to panel **(a)** but showing the Pearson correlation between the NN-predicted summer surface nitrate concentration and the summer surface total chlorophyll concentration from the Copernicus reanalysis. Panel **(d)** is the same as panel **(b)** but showing interannual variability of NN-predicted surface nitrate concentration in the summer rather than in the winter. The dashed contours in panels **(a)** and **(c)** show regions where the correlation is statistically significant (p value < 0.05).

NWES nitrate reanalysis. Using the newly developed product, we have identified nutrient-limited coastal areas with potentially strong ecosystem responses to an event of river nutrient pollution, addressed nitrate decadal trends, and tested how successfully winter nitrate can be used as a predictor of the phytoplankton spring bloom. The areas of strong ecosystem response to nutrient loads were identified as the southeast of the North Sea, the coastline of Brittany, and the Bay of Biscay (areas previously marked by OSPAR as eutrophication problem areas, Axe et al., 2017; Devlin et al., 2023), but additional coastal areas were also identified in the south of Ireland, the eastern Scottish coastline, and the Irish Sea. We have found that nitrate trends in the last 2 decades were mostly minor, with the exception of the coastline in the Bay of Biscay. This confirms recent observations by Axe et al. (2017, 2022) and Devlin et al. (2023). We have also demonstrated that winter nitrate is only a limited predictor of the

next season's growth, which again supports recent findings of Van Leeuwen et al. (2023).

There are many other potential scientific uses for the nitrate dataset; e.g. we propose assimilating the nitrate data into the NWES operational model to correct the model's significant nitrate biases, potentially improving its dynamics and its short-range forecasts. The model's skill in simulating phytoplankton is known to quickly degrade with the forecast lead time (e.g. Kay et al., 2016; Skákala et al., 2018), and biases in nitrate might be one of the leading factors in driving this. Assimilation of nitrate products derived here will be addressed in the near future.

Several extensions of this work would also be desirable, such as utilising ICES data for other biogeochemical indicators to produce ML-informed multi-variate datasets across the whole NWES domain (these should include other nutrients and oxygen). ML could also identify valuable patterns
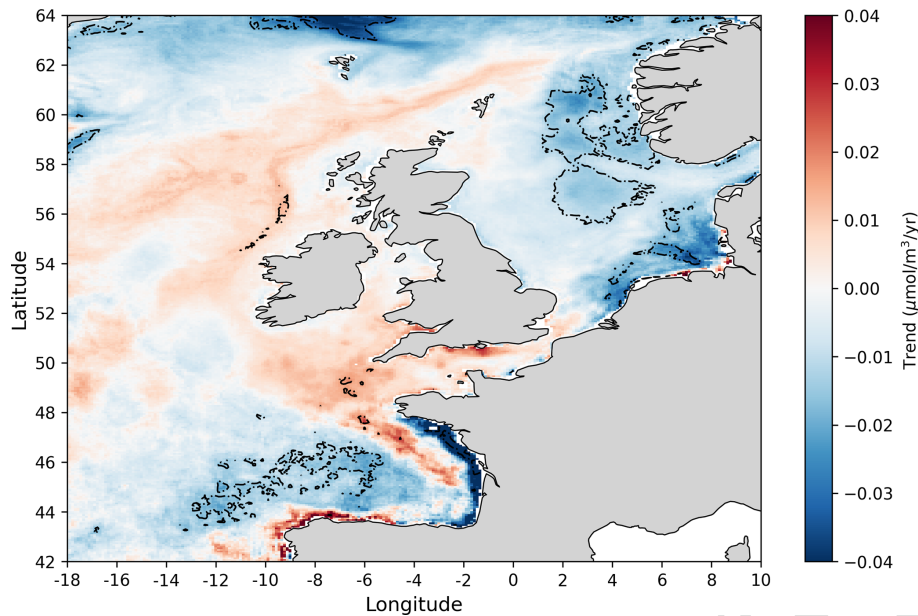
Figure 7. Linear trends at each spatial location in the annual nitrate concentration 1998–2020 time series obtained from the NN model prediction. Dashed contours mark areas with statistically significant trends (p value < 0.05).

of relationships across the multiple variables. Furthermore, the model developed here did not show very good skill in capturing high-frequency (daily) temporal variability, including extreme events. This might be due to processes provid-
⁵ ing the ocean with a memory that is significantly longer than the daily timescale of the product. Representing ocean memory by the NN model might require using time-lagged input features, which could substantially inflate the size and the complexity of the model. Despite this, including such
¹⁰ features into the NN model should be considered in the future, i.e. also with respect to developing a more realistic river discharge advection scheme than the one used here, which is a major challenge. Interestingly, the finer spatio-temporal representation of nitrate could be automatically im-
¹⁵ proved if we assimilated the NN-predicted nitrate into the dynamical model, as such a scheme would benefit from both the dynamical model advection scheme and the improved representation of the (coarser-resolution) nitrate by the NN model. Other future activities should include re-training the
²⁰ NN model whenever newer and potentially better observational products appear for its inputs (such as new riverine discharge data appearing in the last few years, van Leeuwen and Lenhart, 2021). Finally, ML tools designed to specifically capture extreme phenomena can be deployed in the fu-
²⁵ ture and extend the applicability of this work.

Code and data availability. The ML software is available at https://github.com/neccton-algo/nn-bgc (Banerjee, 2025). In this study we used the atmospheric ERA5 product of the European Centre for Medium-Range Weather Forecasts (ECMWF;

³⁰ https://doi.org/10.24381/cds.143582cf, Hersbach et al., 2017) and river data that are stored on the MonSOON HPC, which can be obtained upon request. We also used EU Copernicus reanalyses, the NWSHELF_MULTIYEAR_BGC_004_011 (https://doi.org/10.48670/moi-00059) product for biogeochem-
³⁵ istry, and the NWSHELF_MULTIYEAR-_PHY-_004_009 (https://doi.org/10.48670/moi-00058, Ciavatta et al., 2018) product for physics. We used nitrate datasets from the ICES portal (https://doi.org/10.17895/ices.pub.8883), the Western Channel Observatory (https://www.westernchannelobservatory.org.uk/
l4_nutrients.php, Western Channel Observatory, 2025), and the ⁴⁰ Scottish Coastal Observatory (Scottish Coastal Observatory, 2018a, https://doi.org/10.7489/12138-1; Scottish Coastal Observatory, 2018b https://doi.org/10.7489/610-1; Scottish Coastal Observatory, 2018c, https://doi.org/10.7489/948-1; Scottish Coastal Observatory, 2018d, https://doi.org/10.7489/952-1; Scottish Coastal ⁴⁵ Observatory, 2018e, https://doi.org/10.7489/953-1).

Supplement. The supplement related to this article is available online at [the link will be implemented upon publication].

Author contributions. DSB wrote all the code, processed the data, developed the ML model, and ran the experiments. JS conceptu- ⁵⁰ alised and supervised the work and acquired funding. Both authors wrote the paper.

Competing interests. The contact author has declared that neither of the authors has any competing interests.

ther geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. This work was funded by the Horizon Europe project titled New Copernicus Capability for Tropic Ocean Networks (NECCTON; grant agreement no. 101081273). We also acknowledge support from the UK Natural Environment Research Council (NERC) National Capability – Science Single Centre Research programme and the Climate Linked Atlantic Sector Science (CLASS) project (NE/R015953/1). The river data used here were prepared by Sonja van Leeuwen and Helen Powley as part of the UK Shelf Sea Biogeochemistry programme (contract no. NE/K001876/1) of the NERC and the Department for Environment, Food and Rural Affairs (DEFRA). The riverine data also contained climatological values from the Global River Discharge Database and the Centre for Ecology & Hydrology (Young and Holt, 2007). We would like to thank Bee Berx for directing us to the validation data from the Scottish coastal stations. We would also like to thank Jerry Blackford, Gennadi Lessin, Yuri Artioli, Helen Powley, Julien Brajard, and Dave Moffat for their valuable comments and discussions.

Financial support. This research has been supported by the Horizon Europe project titled New Copernicus Capability for Tropic Ocean Networks (NECCTON; grant agreement no. 101081273) and the UK Natural Environment Research Council (NERC) National Capability – Science Single Centre Research programme and the Climate Linked Atlantic Sector Science (CLASS) project (NE/R015953/1).

Review statement. This paper was edited by Jamie Shutler and reviewed by two anonymous referees.

References

Anderson, D. M., Cembella, A. D., and Hallegraeff, G. M.: Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management, Annu. Rev. Mar. Sci., 4, 143–176, 2012.

Arrigo, R., Baker, A. R., Capone, D. G., Cornell, S., Dentener, F., Galloway, J., Ganeshram, R. S., Geider, R. J., Jickells, T., Kuypers, M. M., Langlois, R., Liss, P. S., Liu, S. M., Middelburg, J. J., Moore, C. M., Nickovic, S., Oschlies, A., Pedersen, T., Prospero, J., Schlitzer, R., Seitzinger, S., Sorensen, L. L., Uematsu, M., Ulloa, O., Voss, M., Ward, B., and Zamora, L.: Impacts of atmospheric anthropogenic nitrogen on the open ocean, Science, 320, 893–897, 2008.

Axe, P., Clausen, U., Leujak, W., Malcolm, S., and Harvey, E.: Eutrophication status of the OSPAR maritime area, Third Integrated Report on the Eutrophication Status of the OSPAR Maritime Area, https://www.vliz.be/imisdocs/publications/308897.pdf (last access: 20 June 2025), 2017.

Axe, P., Sonesten, L., and Skarbövik, E.: Inputs of Nutrients to the OSPAR Maritime Area, https://oap-cloudfront.ospar.org/media/filer_public/5d/10/5d10879f-0a4e-4c9a-b8b8-138b4215cd0b/p00930_inputs_of_nutrients_qsr2023.pdf (last access: 20 June 2025), 2022.

Banerjee, D. S.: Neural Network Model code, GitHub [code], https://github.com/dsbanerjee90/neccton_algo_bgcnn, last access: 20 June 2025.

Baretta, J., Ebenhöh, W., and Ruardij, P.: The European regional seas ecosystem model, a complex marine ecosystem model, Neth. J. Sea Res., 33, 233–246, 1995.

Behrenfeld, M. J. and Boss, E. S.: Resurrecting the ecological underpinnings of ocean plankton blooms, Annu. Rev. Mar. Sci., 6, 167–194, 2014.

Beman, J. M., Popp, B. N., and Francis, C. A.: Molecular and biogeochemical evidence for ammonia oxidation by marine Crenarchaeota in the Gulf of California, ISME J., 2, 429–441, 2008.

Board, O. S. and Council, N. R.: Clean coastal waters: understanding and reducing the effects of nutrient pollution, National Academies Press, ISBN 9780309069489, 2000.

Borges, A., Schiettecatte, L.-S., Abril, G., Delille, B., and Gazeau, F.: Carbon dioxide in European coastal waters, Estuar. Coast. Shelf S., 70, 375–387, 2006.

Bresnan, E., Cook, K., Hindson, J., Hughes, S., Lacaze, J., Walsham, P., Webster, L., and Turrell, W.: The Scottish coastal observatory 1997–2013. Part 2-description of Scotland's coastal waters, Scottish Marine and Freshwater Science, 7, https://marine.gov.scot/sma/content/scottish-coastal-observatory-1997-2013-part-2-description-scotlands-coastal-waters (last access: 20 June 2025), 2016.

Brewin, R. J., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., and Hardman-Mountford, N. J.: A three-component model of phytoplankton size class for the Atlantic Ocean, Ecol. Model., 221, 1472–1483, 2010.

Brewin, R. J., Ciavatta, S., Sathyendranath, S., Jackson, T., Tilstone, G., Curran, K., Airs, R. L., Cummings, D., Brotas, V., Organelli, E., Dall'Olmo, G., and Dionysios, R. E.: Uncertainty in ocean-color estimates of chlorophyll for phytoplankton groups, Frontiers in Marine Science, 4, 104, https://doi.org/10.3389/fmars.2017.00104, 2017.

Brockmann, U. and Eberlein, K.: River input of nutrients into the German Bight, in: The Role of Freshwater Outflow in Coastal Marine Ecosystems, Springer, 231–240, https://doi.org/10.1007/978-3-642-70886-2_15, 1986.

Brockmann, U., Topcu, D., Schütt, M., and Leujak, W.: Eutrophication assessment in the transit area German Bight (North Sea) 2006–2014–stagnation and limitations, Mar. Pollut. Bull., 136, 68–78, 2018.

Bruggeman, J. and Bolding, K.: A general framework for aquatic biogeochemical models, Environ. Modell. Softw., 61, 249–265, 2014.

Burson, A., Stomp, M., Akil, L., Brussaard, C. P., and Huisman, J.: Unbalanced reduction of nutrient loads has created an offshore gradient from phosphorus to nitrogen limitation in the North Sea, Limnol. Oceanogr., 61, 869–888, https://doi.org/10.1002/LNO.10257, 2016.

Butenschön, M., Clark, J., Aldridge, J. N., Allen, J. I., Artioli, Y., Blackford, J., Bruggeman, J., Cazenave, P., Ciavatta, S., Kay, S., Lessin, G., van Leeuwen, S., van der Molen, J., de Mora, L., Polimene, L., Sailley, S., Stephens, N., and Torres, R.: ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels, Geosci. Model Dev., 9, 1293–1339, https://doi.org/10.5194/gmd-9-1293-2016, 2016.

Chen, S., Meng, Y., Lin, S., Yu, Y., and Xi, J.: Estimation of sea surface nitrate from space: current status and future potential, Sci. Total Environ., 899, 165690, https://doi.org/10.1016/j.scitotenv.2023.165690, 2023.

Ciavatta, S., Brewin, R. J. W., Skákala, J., Polimene, L., de Mora, L., Artioli, Y., and Allen, J. I.: Assimilation of ocean-color plankton functional types to improve marine ecosystem simulations, J. Geophys. Res.-Oceans, 123, 834–854, https://doi.org/10.1002/2017JC013490, 2018 (data available at: https://doi.org/10.48670/moi-00058).

Devlin, M. J., Prins, T. C., Enserink, L., Leujak, W., Heyden, B., Axe, P. G., Ruiter, H., Blauw, A., Bresnan, E., Collingridge, K., Devreker, D., Fernand, L., Jakobsen, Gomez J, F., Graves, C., Lefebvre, A., Lenhart, H., Markager, S., Nogueira, M., O'Donnell, G., Parner, H., Skarbovik, E., Skogen, D. Morten, S. L., Van Leeuwen, S. M., Wilkes, R., Dening, E., and Iglesias-Campos, A.: A first ecological coherent assessment of eutrophication across the North-East Atlantic waters (2015–2020), Frontiers in Ocean Sustainability, 1, 1253923, https://doi.org/10.3389/focsu.2023.1253923, 2023.

Diaz, R. J. and Rosenberg, R.: Spreading dead zones and consequences for marine ecosystems, Science, 321, 926–929, 2008.

Doney, S. C., Fabry, V. J., Feely, R. A., and Kleypas, J. A.: Ocean acidification: the other CO_2 problem, Annu. Rev. Mar. Sci., 1, 169–192, 2009.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The operational sea surface temperature and sea ice analysis (OSTIA) system, Remote Sens. Environ., 116, 140–158, 2012.

Duarte, C. M.: Coastal eutrophication research: a new awareness, in: Eutrophication in Coastal Ecosystems: Towards better understanding and management strategies Selected Papers from the Second International Symposium on Research and Management of Eutrophication in Coastal Ecosystems, 20–23 June 2006, Nyborg, Denmark, Springer, 263–269, https://doi.org/10.1007/s10750-009-9795-8, 2009.

Durairaj, P., Sarangi, R. K., Ramalingam, S., Thirunavukarassu, T., and Chauhan, P.: Seasonal nitrate algorithms for nitrate retrieval using OCEANSAT-2 and MODIS-AQUA satellite data, Environ. Monit. Assess., 187, 1–15, 2015.

Dutkiewicz, S., Follows, M., Marshall, J., and Gregg, W. W.: Interannual variability of phytoplankton abundances in the North Atlantic, Deep-Sea Res. Pt. II, 48, 2323–2344, 2001.

Follows, M. and Dutkiewicz, S.: Meteorological modulation of the North Atlantic spring bloom, Deep-Sea Res. Pt. II, 49, 321–344, 2001.

Garcia, H., Weathers, K., Paver, C., Smolyar, I., Boyer, T., Locarnini, M., Zweng, M., Mishonov, A., Baranova, O., Seidov, D., Reagan, J., 375 Paver, C., Smolyar, I., Boyer, T., Locarnini, M., Zweng, M., Mishonov, A., Baranova, O., Seidov, D., and Reagan, J.: World Ocean Atlas 2018, vol. 4, Dissolved inorganic nutrients (phosphate, nitrate and nitrate+nitrite, silicate), https://archimer.ifremer.fr/doc/00651/76336/ (last access: 20 June 2025), 2019.

Good, S., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Reid, R., Roberts-Jones, J., Searle, T., Waters, J., While, J., and Worsfold, M.: The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses, Remote Sens.-Basel, 12, 720, https://doi.org/10.3390/rs12040720, 2020.

Greenwood, N., Devlin, M. J., Best, M., Fronkova, L., Graves, C. A., Milligan, A., Barry, J., and Van Leeuwen, S. M.: Utilizing eutrophication assessment directives from transitional to marine systems in the Thames Estuary and Liverpool Bay, UK, Frontiers in Marine Science, 6, 116, https://doi.org/10.3389/fmars.2019.00116, 2019.

Grosse, J., van Breugel, P., Brussaard, C. P., and Boschker, H. T.: A biosynthesis view on nutrient stress in coastal phytoplankton, Limnol. Oceanogr., 62, 490–506, 2017.

Harris, R.: The L4 time-series: the first 20 years, J. Plankton Res., 32, 577–583, 2010.

He, R., Chen, K., Fennel, K., Gawarkiewicz, G. G., and McGillicuddy Jr, D. J.: Seasonal and interannual variability of physical and biological dynamics at the shelfbreak front of the Middle Atlantic Bight: nutrient supply mechanisms, Biogeosciences, 8, 2935–2946, https://doi.org/10.5194/bg-8-2935-2011, 2011.

Henson, S. A., Robinson, I., Allen, J. T., and Waniek, J. J.: Effect of meteorological conditions on interannual variability in timing and magnitude of the spring bloom in the Irminger Basin, North Atlantic, Deep-Sea Res. Pt. I, 53, 1601–1615, 2006.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: Complete ERA5 from 1940: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus Climate Change Service (C3S) Data Store (CDS) [data set], https://doi.org/10.24381/cds.143582cf, 2017.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/QJ.3803, 2020.

Hill, R., Rinker, R., and Wilson, H. D.: Atmospheric nitrogen fixation by lightning, J. Atmos. Sci., 37, 179–192, 1980.

Hindson, J., Berx, B., Hughes, S., Walsham, P., Machairopoulou, M., Bresnan, E., and Turrell, B.: The Scottish Coastal Observatory, Bollettino di Geofisica, 333, https://www.vliz.be/imisdocs/

publications/321529.pdf#page=335 (last access: 20 June 2025), 2018.

Hinrichs, I., Gouretski, V., Pätch, J., Emeis, K., and Stammer, D.: North sea biogeochemical climatology, https://pure.mpg.de/rest/items/item_2478691/component/file_3185079/content (last access: 20 June 2025), 2017.

Huisman, J., van Oostveen, P., and Weissing, F. J.: Critical depth and critical turbulence: two different mechanisms for the development of phytoplankton blooms, Limnol. Oceanogr., 44, 1781–1787, 1999.

Huthnance, J. M., Holt, J. T., and Wakelin, S. L.: Deep ocean exchange with west-European shelf seas, Ocean Sci., 5, 621–634, 2009.

Jahnke, R. A.: Global synthesis, in: Carbon and nutrient fluxes in continental margins: A global synthesis, Springer, 597–615, ISBN 978-3-662-51787-1, https://doi.org/10.1007/978-3-540-92735-8_16, 2010.

Jin, H., Song, Q., and Hu, X.: Auto-Keras: An Efficient Neural Architecture Search System, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19, Association for Computing Machinery, 1946–1956, ISBN 9781450362016, https://doi.org/10.1145/3292500.3330648, 2019.

Kay, S., McEwan, R., and Ford, D.: North West European Shelf Production Centre NWSHELF_MULTIYEAR_BIO_004_011, CMEMS Report, 3, 21, https://doi.org/10.48670/moi-00058, 2016.

Lenhart, H.-J. and Große, F.: Assessing the effects of WFD nutrient reductions within an OSPAR frame using transboundary nutrient modeling, Frontiers in Marine Science, 5, 447, https://doi.org/10.3389/fmars.2018.00447, 2018.

Lenhart, H.-J., Mills, D. K., Baretta-Bekker, H., van Leeuwen, S. M., van der Molen, J., Baretta, J. W., Blaas, M., Desmit, X., Kühn, W., Lacroix, G., Los, H. J., Ménesguen, A., Neves, R., Proctor, R., Ruardij, P., Skogen, M. D., Vanhoutte-Brunier, A., Villars, M. T., and Wakelin, S. L.: Predicting the consequences of nutrient reduction on the eutrophication status of the North Sea, J. Marine Syst., 81, 148–170, https://doi.org/10.1016/j.jmarsys.2009.12.014, 2010.

Loebl, M., Colijn, F., van Beusekom, J. E., Baretta-Bekker, J. G., Lancelot, C., Philippart, C. J., Rousseau, V., and Wiltshire, K. H.: Recent patterns in potential phytoplankton limitation along the Northwest European continental coast, J. Sea Res., 61, 34–43, 2009.

Madec, G., Bourdallé-Badie, R., Bouttier, P.-A., Bricaud, C., Bruciaferri, D., Calvert, D., Chanut, J., Clementi, E., Coward, A., and Delrosso, D.: NEMO Ocean Engine, Zenodo [data set], https://doi.org/10.5281/zenodo.3248739, 2017.

Nazari-Sharabian, M., Ahmad, S., and Karakouzian, M.: Climate change and eutrophication: a short review, Engineering, Technology and Applied Science Research, 8, 3668, https://digitalscholarship.unlv.edu/fac_articles/562 (last access: 20 June 2025), 2018.

Noxon, J.: Atmospheric nitrogen fixation by lightning, Geophys. Res. Lett., 3, 463–465, 1976.

OSPAR Commission: Common Procedure for the Identification of the Eutrophication Status of the OSPAR Maritime Area, OSPAR Commission, 3, https://oap.ospar.org/en/ospar-assessments/intermediate-assessment-2017/pressures-human-activities/

eutrophication/third-comp-summary-eutrophication/ (last access: 20 June 2025), 2005.

Painting, S., Van der Molen, J., Parker, E., Coughlan, C., Birchenough, S., Bolam, S., Aldridge, J., Forster, R., and Greenwood, N.: Development of indicators of ecosystem functioning in a temperate shelf sea: a combined fieldwork and modelling approach, Biogeochemistry, 113, 237–257, 2013.

Pauly, D., Christensen, V., Guénette, S., Pitcher, T. J., Sumaila, U. R., Walters, C. J., Watson, R., and Zeller, D.: Towards sustainability in world fisheries, Nature, 418, 689–695, 2002.

Philippart, C. J., Beukema, J. J., Cadée, G. C., Dekker, R., Goedhart, P. W., van Iperen, J. M., Leopold, M. F., and Herman, P. M.: Impacts of nutrient reduction on coastal communities, Ecosystems, 10, 96–119, 2007.

Postgate, J. R.: Nitrogen Fixation, Cambridge University Press, ISBN 9780521648530, 1998.

Rabalais, N. N., Turner, R. E., and Wiseman Jr., W. J.: Gulf of Mexico hypoxia, aka "The dead zone", Annu. Rev. Ecol. Syst., 33, 235–263, 2002.

Rabalais, N. N., Turner, R. E., Díaz, R. J., and Justić, D.: Global change and eutrophication of coastal waters, ICES J. Mar. Sci., 66, 1528–1537, 2009.

Radach, G.: Ecosystem functioning in the German Bight under continental nutrient inputs by rivers, Estuaries, 15, 477–496, 1992.

Radach, G. and Pätsch, J.: Variability of continental riverine freshwater and nutrient inputs into the North Sea for the years 1977–2000 and its consequences for the assessment of eutrophication, Estuar. Coast., 30, 66–81, 2007.

Ryther, J. H. and Dunstan, W. M.: Nitrogen, phosphorus, and eutrophication in the coastal marine environment, Science, 171, 1008–1013, 1971.

Scottish Coastal Observatory: Dataset – 12138-1: Scottish Coastal Observatory Dataset – 12138-1, [data set] https://doi.org/10.7489/12138-1, 2018a.

Scottish Coastal Observatory: Dataset – 610-1: Scottish Coastal Observatory Dataset – 610-1, [data set], https://doi.org/10.7489/610-1, 2018b.

Scottish Coastal Observatory: Dataset – 948-1: Scottish Coastal Observatory Dataset – 948-1, [data set], https://doi.org/10.7489/948-1, 2018c.

Scottish Coastal Observatory: Dataset – 952-1: Scottish Coastal Observatory Dataset - 952-1, [data set], https://doi.org/10.7489/952-1, 2018d.

Scottish Coastal Observatory: Dataset – 953-1: Scottish Coastal Observatory Dataset – 953-1, [data set], https://doi.org/10.7489/953-1, 2018e.

Skákala, J., Ford, D., Brewin, R. J., McEwan, R., Kay, S., Taylor, B., de Mora, L., and Ciavatta, S.: The assimilation of phytoplankton functional types for operational forecasting in the northwest European shelf, J. Geophys. Res.-Oceans, 123, 5230–5247, 2018.

Skákala, J., Ford, D., Bruggeman, J., Hull, T., Kaiser, J., King, R. R., Loveday, B., Palmer, M. R., Smyth, T., Williams, C. A., and Ciavatta, S.: Towards a multi-platform assimilative system for North Sea biogeochemistry, J. Geophys. Res.-Oceans, 126, e2020JC016649, https://doi.org/10.1029/2020JC016649, 2021.

Skákala, J., Bruggeman, J., Ford, D., Wakelin, S., Akpınar, A., Hull, T., Kaiser, J., Loveday, B. R., O'Dea, E., Williams, C. A., and Ciavatta, S.: The impact of ocean biogeochemistry on

physics and its consequences for modelling shelf seas, Ocean Model., 172, 101976, https://doi.org/10.1029/2018JC014153, 2022.

Skogen, M. D., Søiland, H., and Svendsen, E.: Effects of changing nutrient loads to the North Sea, J. Marine Syst., 46, 23–38, 2004.

Soetaert, K., Middelburg, J. J., Heip, C., Meire, P., Van Damme, S., and Maris, T.: Long-term change in dissolved inorganic nutrients in the heterotrophic Scheldt estuary (Belgium, The Netherlands), Limnol. Oceanogr., 51, 409–423, 2006.

Sonesten, L., Axe, P., Bellert, B., Burtschell, L., Eumont, D., Fairbank, V., Farkas, C., Graves, C., Martínez García-Denche, L., McDermott, G., Moeslund Svendsen, L., Mönnich, J., Nunes, S., Pohl, M., Posen, P., Sánchez Fernández, B., Skarbøvik, E., Thiesse, E., Vannevel, R., and Wilkes, R.: Waterborne and Atmospheric Inputs of Nutrients and Metals to the Sea, in: OSPAR, 2023: The 2023 Quality Status Report for the Northeast Atlantic, OSPAR Commission, London, https://oap.ospar.org/en/ospar-assessments/quality-status-reports/qsr-2023/other-assessments/inputs-nutrients-and-metals (last access: 20 June 2025), 2022.

Sverdrup, H. U.: On conditions for the vernal blooming of phytoplankton, J. Cons. Int. Explor. Mer, 18, 287–295, 1953.

Tett, P., Droop, M. R., and Heaney, S. I.: The redfield ratio and phytoplankton growth rate, J. Mar. Biol. Assoc. UK, 65, 487–504, https://doi.org/10.1017/S0025315400050566, 1985.

Topcu, D. and Brockmann, U.: Consistency of thresholds for eutrophication assessments, examples and recommendations, Environ. Monit. Assess., 193, 1–15, 2021.

Ueyama, R. and Monger, B. C.: Wind-induced modulation of seasonal phytoplankton blooms in the North Atlantic derived from satellite observations, Limnol. Oceanogr., 50, 1820–1829, 2005.

van Leeuwen, S. and Lenhart, H.: OSPAR ICG-EMO riverine database 2020-05-01 used in 2020 workshop, NIOZ, the Royal Netherlands Institute for Sea Research [data set], https://doi.org/10.25850/nioz/7b.b.vc, 2021.

Van Leeuwen, S. M., Lenhart, H.-J., Prins, T. C., Blauw, A., Desmit, X., Fernand, L., Friedland, R., Kerimoglu, O., Lacroix, G., Van Der Linden, A., Lefebvre, A., Molen, J. v. d., Plus, M., Baroni, I. R., Silva, T., Stegert, C., Troost, T. A., and Vilmin, L.: Deriving pre-eutrophic conditions from an ensemble model approach for the North-West European seas, Frontiers in Marine Science, 10, 1129951, https://doi.org/10.3389/fmars.2023.1129951, 2023.

Voss, M., Bange, H. W., Dippner, J. W., Middelburg, J. J., Montoya, J. P., and Ward, B.: The marine nitrogen cycle: recent discoveries, uncertainties and the potential relevance of climate change, Philos. T. Roy. Soc. B, 368, 20130121, https://doi.org/10.1098/rstb.2013.0121, 2013.

Western Channel Observatory: L4 Nutrients – Time Series, https://www.westernchannelobservatory.org.uk/l4_nutrients.php, last access: 20 June 2025.

Withers, P. J., Neal, C., Jarvie, H. P., and Doody, D. G.: Agriculture and eutrophication: where do we go from here?, Sustainability, 6, 5853–5875, 2014.

Young, E. and Holt, J.: Prediction and analysis of long-term variability of temperature and salinity in the Irish Sea, J. Geophys. Res.-Oceans, 112, C01008, https://doi.org/10.1029/2005JC003386, 2007.

Yu, X., Chen, S., and Chai, F.: Remote estimation of sea surface nitrate in the California current system from satellite ocean color measurements, IEEE T. Geosci. Remote, 60, 1–17, 2021.

Zhai, L., Platt, T., Tang, C., Sathyendranath, S., and Walne, A.: The response of phytoplankton to climate variability associated with the North Atlantic Oscillation, Deep-Sea Res. Pt. II, 93, 159–168, 2013.