PML | Plymouth Marine Laboratory

Listen to the ocean

Combining machine learning and data assimilation to improve phytoplankton forecasting on the North-West European Shelf

Jozef Skakala, Deep S Banerjee (PML, NCEO) and David Ford (Met Office)

NECCTON
NEW COPERNICUS CAPABILITY FOR TROPHIC OCEAN NETWORKS

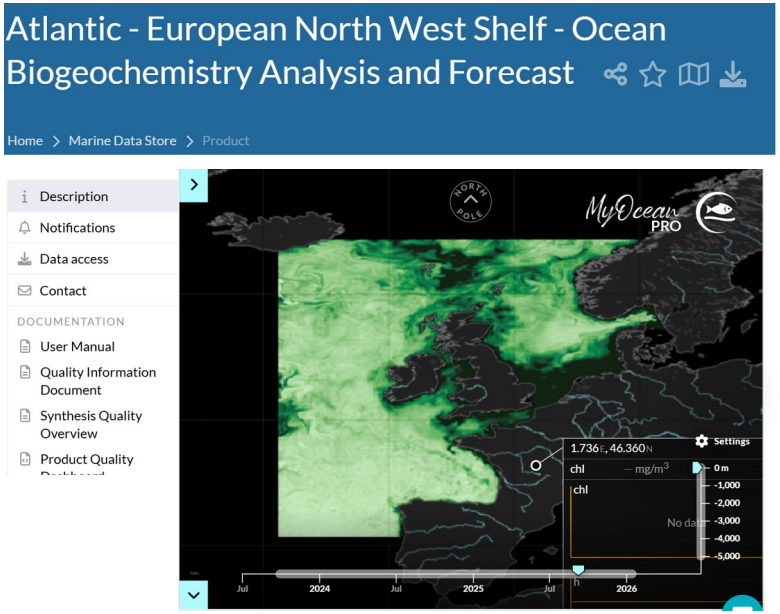
European Commission

UKRI Natural Environment Research Council

Met Office

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

Marine biogeochemistry operational system on the North-West European Shelf



EU Copernicus website

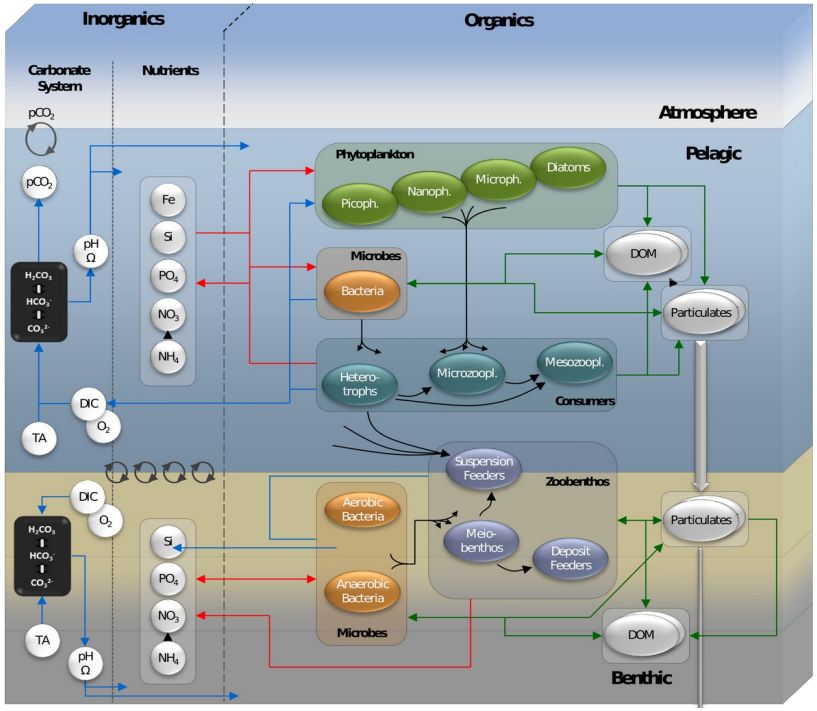
- **NWES is highly productive area of a major interest for European economy, including fisheries and aquaculture**
- **It is impacted by climate and (other) anthropogenic stressors, including marine heatwaves, or eutrophication events..**

Met Office is running an operational system for the NWES, predicting on 5-day time-scale a range of key biogeochemistry indicators, i.e. total/size-class chlorophyll (or phytoplankton carbon biomass), net primary production, oxygen, pH, underwater visibility, nutrients. The predicted variables are supplied to Copernicus, stakeholders such as UK Navy, and were used also to navigate smart observing systems.

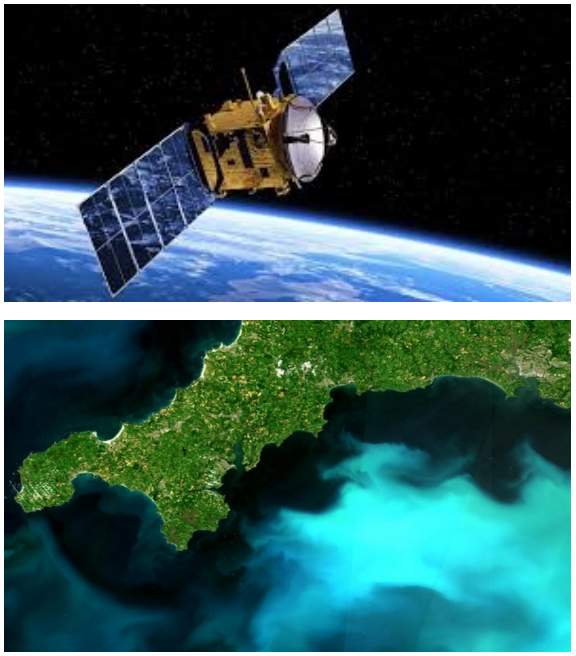
The system design:

The Met Office system runs operationally the NEMO-FABM-ERSEM model on a 7 km resolution scale, with daily assimilation of physics data and satellite ocean-color total chlorophyll concentrations. The assimilation is based on variational (3DVar) NEMOVAR system and updates from marine biogeochemistry phytoplankton biomass Components with other variables updated only through dynamical adjustment.

ERSEM model



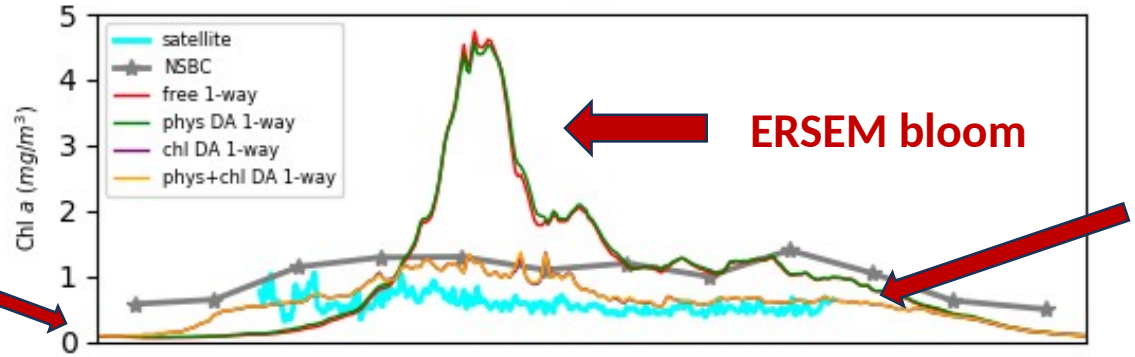
+



Known challenges in the phytoplankton forecasts

ERSEM is characteristic by having a strong Spring bloom that comes relatively late (very late Spring). In the same time it tends to have almost no chlorophyll biomass in the Winter. The reasons for this are likely complex and still not fully understood. The Spring-Autumn chl concentrations in the operational model are corrected via assimilation of ocean color.

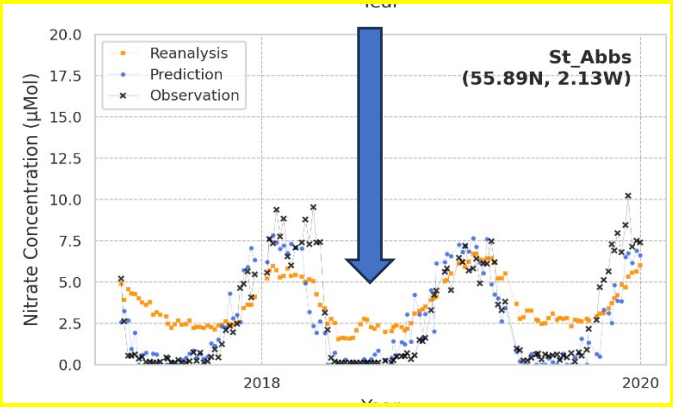
Low Winter concentrations

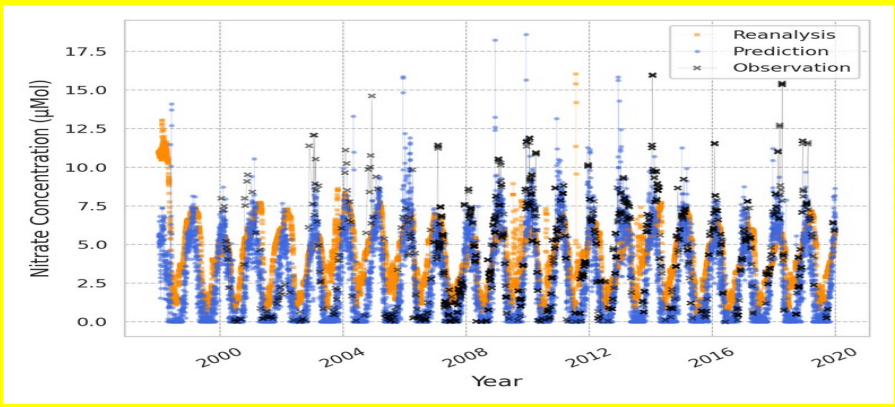


ERSEM bloom

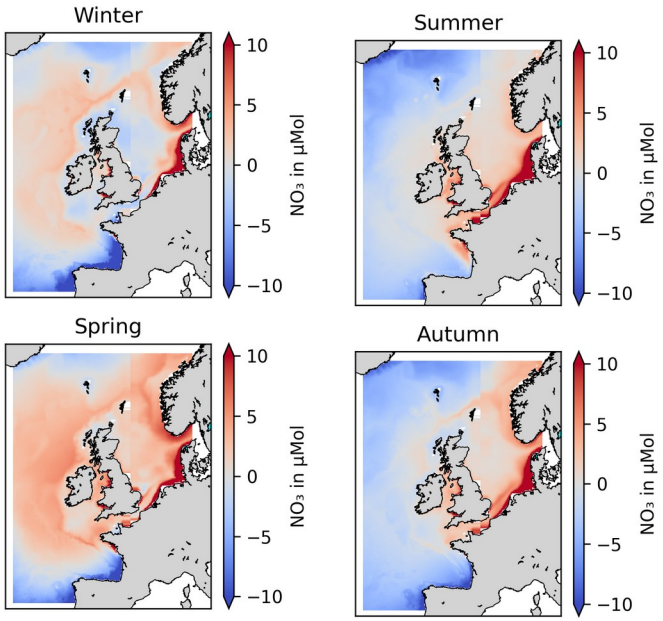
Observations and reanalysis

Limitation of phytoplankton by DA means the Summer nutrients don't get in the analysis depleted, as they would in the nature, or free run.



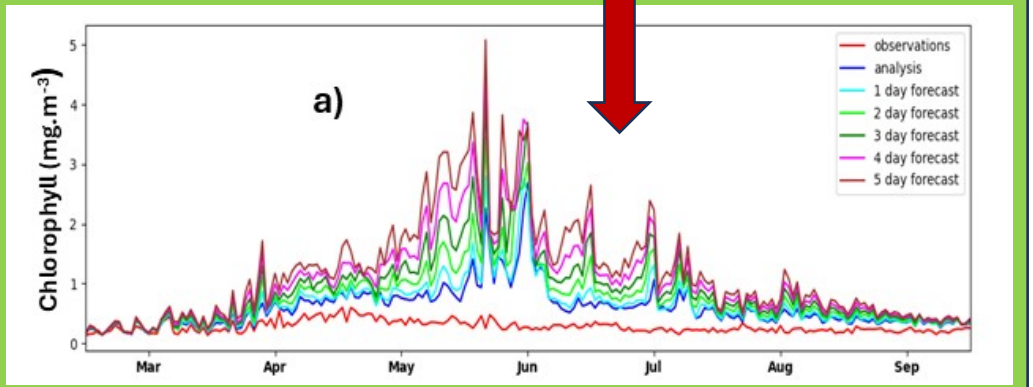


Nitrate degradation in the DA enhances phytoplankton forecast biases



The nitrate (/nutrient) Spring-Summer biases are seen in long-term in the Met Office reanalysis.

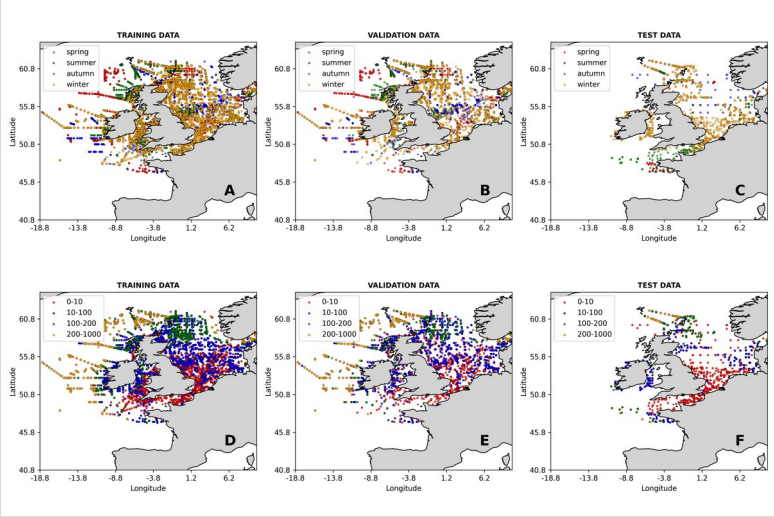
In the late Spring-to-Summer period, there is a rapid growth of phytoplankton chlorophyll forecast bias (NWES-wide) with the forecast lead time. This has the potential to degrade forecast of all the ecosystem indicators. The *Hypothesis is that these biases are caused by the excess nutrients in the DA run.*

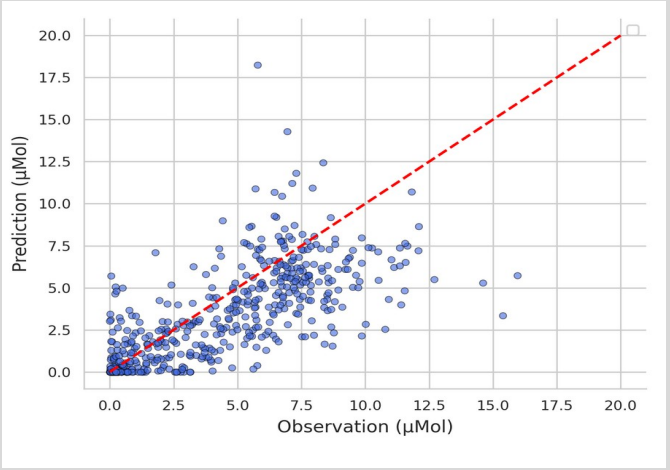


We need to correct nutrients in the DA run, but there is not a simple way how to do that, e.g. proposed balancing schemes have not been so far successful and ensemble methods are too costly and need long-term development… In the same time there is a sparsity of nitrate data to assimilate..

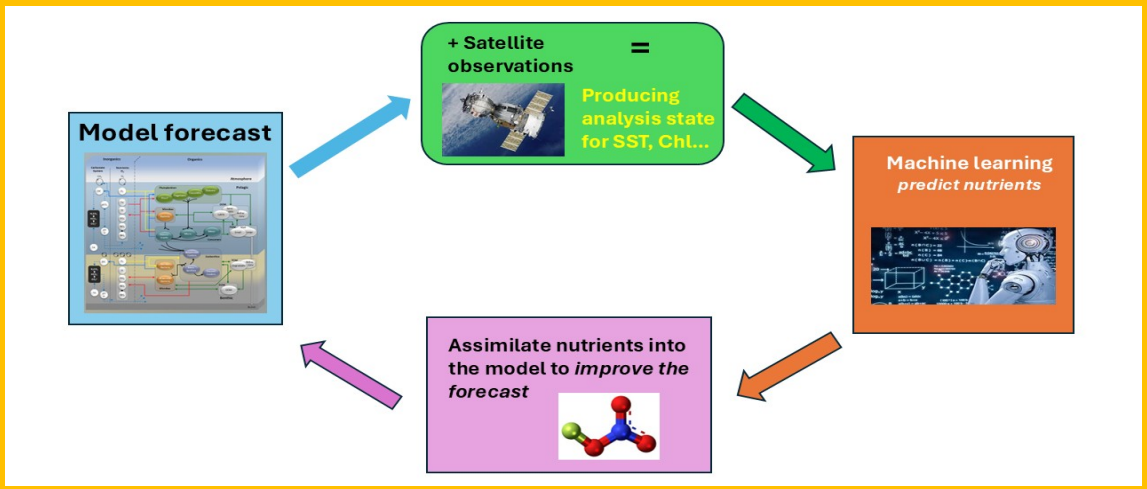
Banerjee and Skakala (2025) developed NN model able to predict nitrate from satellite observations (SST, chlorophyll …) atmospheric data and riverine data. The NN was trained on reanalysis, ERA5, river data-set and ICES nitrate data. It was shown to perform well against independent data-sets and used to produce gap-free surface nitrate across satellite OC period.



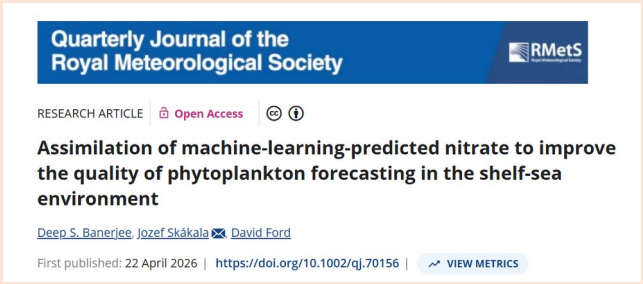


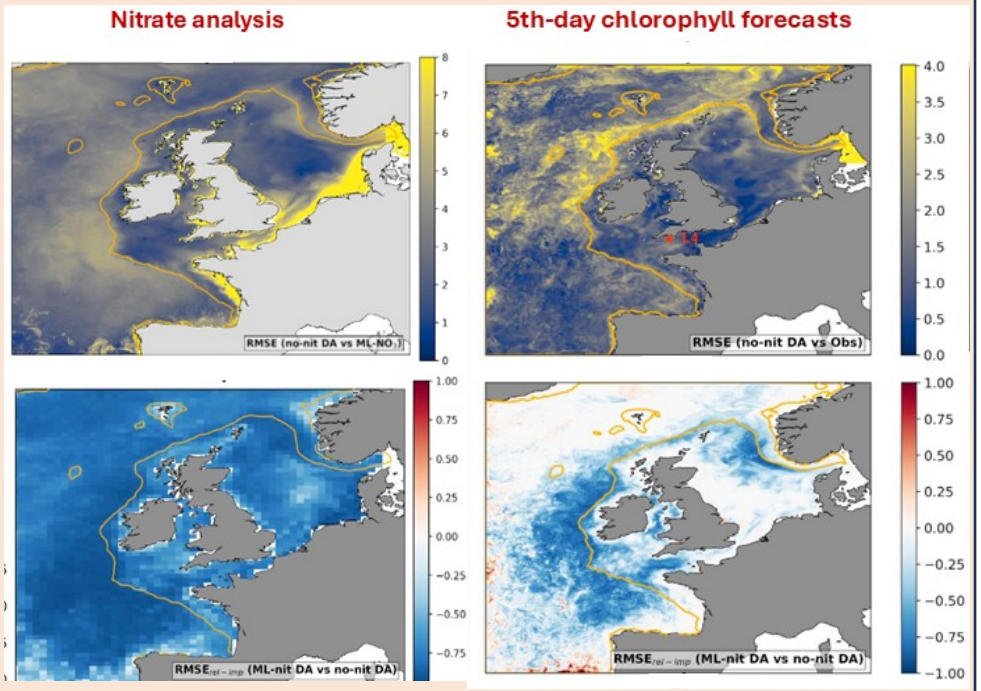


Idea is to deploy the NN model within the operational suite to fix the phytoplankton bias problem

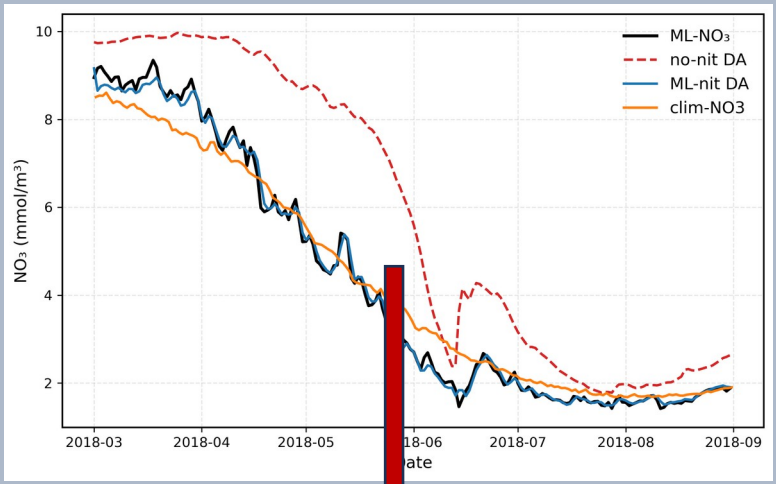


In *Banerjee, Skakala and Ford (2026), published in* QJRMS, we have demonstrated an offline version of this system, i.e. that it substantially improves the phytoplankton forecast skill. We have compared standard forecasting suite with suite assimilating also nitrate.

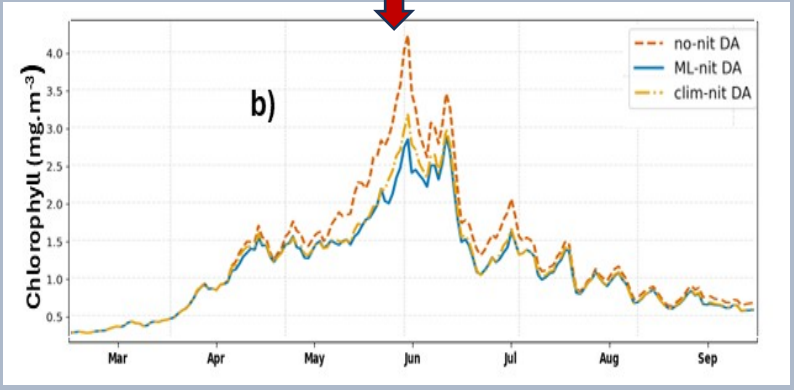

Quarterly Journal of the Royal Meteorological Society
RMetS
RESEARCH ARTICLE | Open Access
Assimilation of machine-learning-predicted nitrate to improve the quality of phytoplankton forecasting in the shelf-sea environment
Deep S. Banerjee, Jozef Skákala, David Ford
First published: 22 April 2026 | https://doi.org/10.1002/qj.70156 | VIEW METRICS



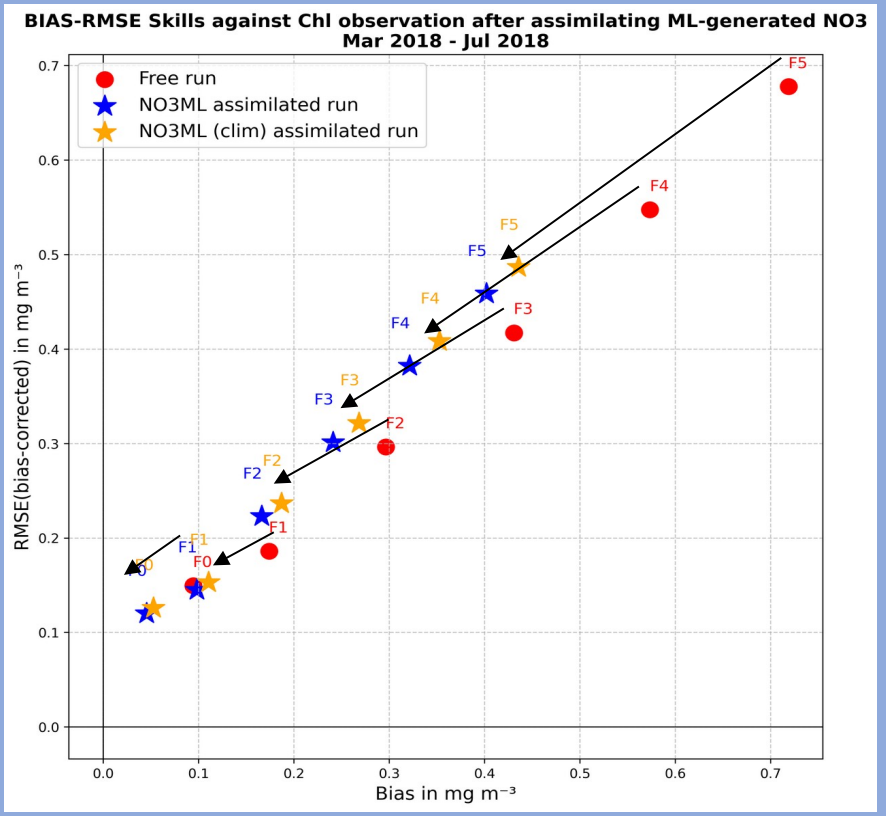
How much do we need flow-dependence?



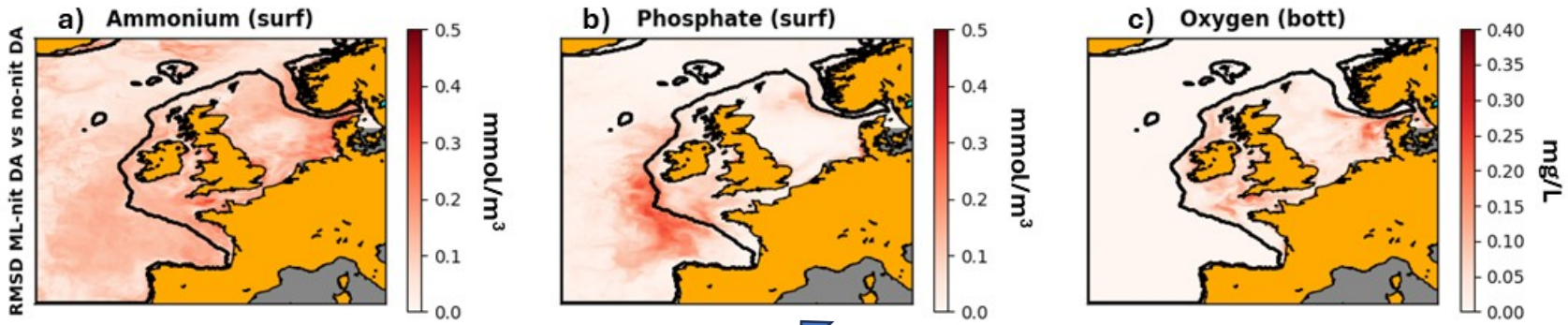
improvement

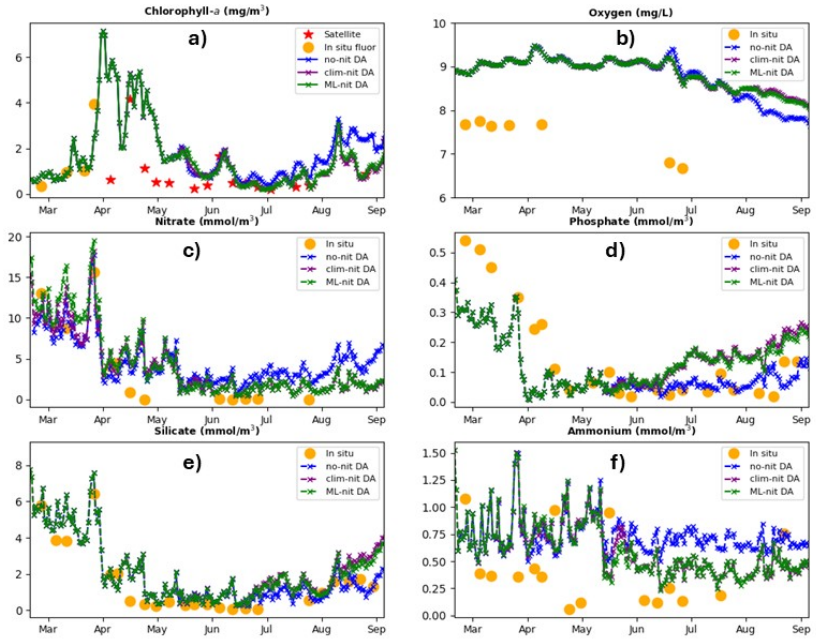


There is consistent improvement (up to 30%) in skill across all the 0-5 lead days. The flow-dependent scheme clearly outperforms the climatological run.



Impact on eutrophication indicators





- NN-derived nitrate assimilation has a clear impact on 5-th day forecast for a range of eutro-phication indicators (beyond nitrate and chlorophyll).
- The impact was validated at the L4 location

Conclusions and future work

- *We have demonstrated that phytoplankton forecast can be substantially improved by combining machine learning with data assimilation, via assimilation of NN predicted nitrate*
- *We have shown that although large improvements are already achieved through assimilating nitrate climatology, using flow-dependent prediction is clearly preferable*
- *The ML-DA system can have significant impact on a range of eutrophication indicators*
- *We plan to implement this system online, with the vision that it becomes an integral part of Met Office operational forecasts. Our analysis shows that the difference between online and offline system should be smaller than the difference between the flow-dependent and climatological nitrate assimilation (and online system is expected to outperform the offline system from this study).*
- *Other future developments are feasible, e.g. increasing the complexity of the ML model, improving the spatio-temporal resolution of predicted nitrate, or predicting other important variables, such as phosphate.*